

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 02:20:31 ; Search time 108 Seconds
(without alignments)
2019.139 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005

Sequence: 1 MTSVSAEHLGKDRPPVLVS.....HLHYELPAADVRRPPIISVRK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2005	100.0	379	2	Q84BC5
2	609.5	30.4	459	2	P94171
3	569	28.4	433	2	Q8GDF1
4	497	24.8	403	2	Q89764
5	425.5	21.2	482	2	Q84GF4
6	424.5	21.2	476	2	Q9RIQ0
7	409	20.4	525	2	Q74211
8	409	20.4	525	2	AA031171
9	374.5	18.7	405	2	Q8SCH7
10	373	18.6	380	2	Q9X3U7
11	360.5	18.0	622	2	Q8GGD7
12	345	17.2	471	2	Q8RIU6
13	342.5	17.1	461	2	Q931L4
14	326	16.3	401	2	Q9R881
15	316	15.8	456	1	REP_STRLI
16	315.5	15.7	312	2	Q8GHP1
17	293.5	14.6	396	2	Q9AGT3
18	277.5	13.8	396	2	Q52220
19	266.5	13.3	528	2	Q57477
20	265	13.2	478	2	Q8RNS3
21	256.5	12.8	505	2	Q83016
22	235.5	11.7	481	2	Q54245
23	230.5	11.5	437	2	Q52205
24	227	11.3	437	1	NEOR_STRCY
25	214.5	10.7	389	2	Q690E5
26	194	9.7	396	2	Q690E9
27	169	8.4	240	2	Q45282
28	163.5	8.2	313	2	Q84FM7
29	161.5	8.1	261	2	Q8GI56
30	127	6.3	314	1	REP_LACHI
31	123.5	6.2	315	2	Q56129

```

32 123 6.1 334 2 Q6DRR7 Q6drr7 streptococc
33 122.5 6.1 315 2 Q842E5 Q842e5 streptococc
34 122.5 6.1 315 2 Q9ENV8 Q9rnv8 streptococc
35 122.5 6.1 315 2 Q9RNM4 Q9rnw4 streptococc
36 121.5 6.1 315 2 Q9RNM5 Q9rnw5 streptococc
37 119 5.9 319 2 Q9ZNC0 Q9znc0 nitrosomona
38 115 5.7 336 2 Q6460 Q6460 synechocyst
39 114.5 5.7 311 2 Q70LG8 Q70lg8 streptococc
40 114.5 5.7 311 2 CAD97600 Cad97600 streptococc
41 114.5 5.7 315 2 Q9X9N4 Q9x9n4 streptococc
42 114 5.7 314 2 Q30850 Q30850 streptococc
43 111.5 5.6 276 2 Q9K4S9 Q9k4s9 pseudomonas
44 111.5 5.6 320 2 Q52769 Q52769 streptococc
45 111.5 5.6 320 2 Q52776 Q52776 streptococc

```

ALIGNMENTS

```

RESULT 1
Q84BC5
ID Q84BC5 PRELIMINARY; PRT; 379 AA.
AC Q84BC5;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Replication protein.
GN Name=rep;
OS Rhodococcus erythropolis.
OG Plasmid pAN12.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22718480; PubMed=12835922;
RA Kostichka K., Tao L., Bramucci M., Tomb J.-F., Nagarajan V., Cheng Q.;
RT "A small cryptic plasmid from Rhodococcus erythropolis:
characterization and utility for gene expression.";
RL Appl. Microbiol. Biotechnol. 62:61-68(2003).
DR EMBL; AY178757; AA031602.1; -
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000989; Rep.
DR Pfam; PF01446; Rep_1; 1.
KW Plasmid.
SQ SEQUENCE 379 AA; 40972 MW; AB53DB0C49736E0D CRC64;

```

```

Query Match 100.0%; Score 2005; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.8e-152;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTSVSAEHLGKDRPPVLVSDDKRGIRHRLPKLQITTTSETFNACGRPISGVNGVTIVN 60
Db 1 MTSVSAEHLGKDRPPVLVSDDKRGIRHRLPKLQITTTSETFNACGRPISGVNGVTIVN 60
Qy 61 GPKSGFGGLRSCGKGIWPCPCAGKGAHRADEISQVAHQLTGSGVAMVTMTMRHTAGQ 120
Db 61 GPKSGFGGLRSCGKGIWPCPCAGKGAHRADEISQVAHQLTGSGVAMVTMTMRHTAGQ 120
Qy 121 RLHDLWTGLSAAKAATNGRRMTEREMYGCDGVRAVEITHGKNGHVVHALLMPSGD 180
Db 121 RLHDLWTGLSAAKAATNGRRMTEREMYGCDGVRAVEITHGKNGHVVHALLMPSGD 180
Qy 181 VSENILLESFSDAMFDRTWSKLVSGLFAAPLRNSGGLDVRKIGGEADQVLAAYLTKIAGSV 240
Db 181 VSENILLESFSDAMFDRTWSKLVSGLFAAPLRNSGGLDVRKIGGEADQVLAAYLTKIAGSV 240
Qy 241 GMEVSGDGKSGRHGNRAPEWIEIADVAVGGDPQALELWREFEFGSGMRRALANSRGLRARA 300
Db 241 GMEVSGDGKSGRHGNRAPEWIEIADVAVGGDPQALELWREFEFGSGMRRALANSRGLRARA 300

```

QY 301 GLGALDTAOIUEEAPVNVALLPARSMMIETCAPYVGEILGLVEAGATWENLRDH 360
DB 301 GLGALDTAOIUEEAPVNVALLPARSMMIETCAPYVGEILGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRPPIISVRK 379
DB 361 LHVRLPAADVRPPIISVRK 379

RESULT 2
P94171
ID P94171 PRELIMINARY; PRT; 459 AA.
AC P94171
DT 01-MAY-1997 (TremBrel. 03, Created)
DT 01-MAY-1997 (TremBrel. 03, Last sequence update)
DT 01-MAR-2004 (TremBrel. 26, Last annotation update)
DE Rep.
GN Name=rep;
OS Arcanobacterium pyogenes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Actinomycetaceae; Actinomycetaceae; Arcanobacterium.
OX NCBI_TaxID=1661;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BBRI;
RX MEDLINE=96292760; PubMed=9620977;
RA Billington S.J., Jost B.H., Songer J.G.;
RT "The Arcanobacterium (Actinomycetes) pyogenes plasmid pAP1 is a member
of the pU101/pJVI family of rolling circle replication plasmids.";
RL J. Bacteriol. 180:3233-3236(1998).
DR EMBL; U83788; AAC46399.1; -.
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000989; Rep.
DR Pfam; PF01446; Rep_1; 1.
KW Plasmid.
SQ SEQUENCE 459 AA; 51264 MW; 2A0A4B0A043BF865 CRC64;

Query Match 30.4%; Score 609.5; DB 2; Length 459;
Best Local Similarity 36.3%; Pred. No. 2.1e-40;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAEHLSGKDRPPVLVSSD-----XRGIRHELRLPKLQIITSETFNACGR-PISGVNCGV 56
DB 67 SATPLGNTVLTFFPVSNESKTKAGRSRERYELRDGLAISTIESVRKCGRPVAPLVSL 126
QY 57 TIVNGPKSGFGSLRSCGKGLWPCPCAGKVGGAHRADEISQVVAHOLGTG-SVAMVTMTWR 115
DB 127 RAKSGKGAGYGGHLTCGVSACVCSAKIAARRKTDLQVVDHAKVHGNTVSMILTQ 186
QY 116 HTAGQRLHDLTGLSAAMKAAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
DB 187 HHKGQGLKHLMDALSTANNRVTSGRRIEFGKQGLGVVRAVEITHGKNGHVVHALL 246
QY 176 MFGSDVSENILSF-----SDAMFDRWTSKLVSGLGFAAPLNSGG 215
DB 247 ISEKD---PLTSTFVYQKQRRRLPYPEIYMSDFIAERWEAGLAKHGVDFT-LRDSGG 302
QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKGRHGNRAPWEIADVA 266
DB 303 LDWTAKDARAION-----YVSKQSTDAISSEVTLGGFKKARNGNRTFPQLADI 354
QY 267 VG-GDQALELWREFFSGMRRRAIAWSRGLRAGLGAELTDAQIUEEAPVNVALL 325
DB 355 LSLGVDVDDLKWKYEKASFGRRALTWSKGLRDWANLGVESQDEEIA-SREIGDEAIALF 413
QY 326 PARSMMIETCAPYVVG--EILGLVEAGA-----TWNLRDLHLHYRLP 366
DB 414 THDAWRQVR-----FGAELLDDVTESGGRAAAYRWLDPRE-IDWSLP 455

RESULT 3
Q8GDF1
ID Q8GDF1 PRELIMINARY; PRT; 433 AA.
AC Q8GDF1
DT 01-MAR-2003 (TremBrel. 23, Created)
DT 01-MAR-2003 (TremBrel. 23, Last sequence update)
DT 01-MAR-2004 (TremBrel. 26, Last annotation update)
DE Rep.
GN Name=rep;
OS Propionibacterium granulosum.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=33011;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PF283;
RA Farrar M.D., Holland K.T.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY150274; AAN78123.1; -.
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000989; Rep.
DR Pfam; PF01446; Rep_1; 1.
KW Plasmid.
SQ SEQUENCE 433 AA; 48340 MW; 1A8C34C0FB678F CRC64;

Query Match 28.4%; Score 569; DB 2; Length 433;
Best Local Similarity 35.6%; Pred. No. 3.5e-37;
Matches 136; Conservative 63; Mismatches 145; Indels 38; Gaps 14;

QY 16 PVLVSSDKRGIRHELRLPKLQIITSETFNACGR-PISGVNCGVTHGKNGHVVHALLMFSGD---VSENI--- 185
DB 51 PCRKNENGRHRYENRDLRNPQVPLERVRKCGAVPVS--QRIALMAGHGAGYAGLAT 108
QY 73 CGKWCICPCAGKVGGAHRADEISQVVAHOLGTG-SVAMVTMTWRHTAGORLHDLTGLSA 131
DB 109 CGSVWACPVCAAKISAHRRDELARVVQVAVGLGFKVSMILTQRRHAGGDLAELWASLQS 168
QY 132 AWAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALLMFSGD---VSENI--- 185
DB 169 GNAVTSGRWQEFCAQLGVQGVKXAVETHGSHGHVHVHLVISKODPTSVDTKIRHR 228
QY 186 -----LFSFSDAMFDRWTSKLVSGLG--FAAPLNSGGILDKIGGEADQVLA 230
DB 229 RKQGRRTTYPPEVORPEPTAERWSRGLRKGVDVFLA---GSGGLDWQATDSGDEALG 285
QY 231 AYLTKIAS---GVGMEVSGDGKGRHGNRAPWEIADVAVG-GDQALELWREFFSGMG 286
DB 286 RYVAKMNSVDCGLANEATLGGFKKARNGNRTFPQILEDFLDTGSETDLRLWRTYVSASHG 345
QY 287 RRAIWSRGLRAGLGAELTDAQIUEEAPVNVALLIPARSMMIETCAPYVGEI- 344
DB 346 RKALTSKGLRDWAGMESEDEQVAAQDCGEA-VALPFDHAWRQIRTAGAFLDLE 404
QY 345 LGLVEAGATWENLRDLHLHYRLP 366
DB 405 LHSGEGVYAWLKKR-RHIEIP 425

RESULT 4
O69764
ID O69764 PRELIMINARY; PRT; 403 AA.
AC O69764
DT 01-AUG-1998 (TremBrel. 07, Created)
DT 01-AUG-1998 (TremBrel. 07, Last sequence update)
DT 01-MAR-2004 (TremBrel. 26, Last annotation update)
DE Putative rep protein.
GN Name=rep;
OS Propionibacterium freudenreichii.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

DR GO:0006260; P:DNA replication; IEA.
 DR InterPro: IPR000989; Rep.
 DR Pfam: PF01446; Rep_1; 1.
 KW Plasmid.
 SQ SEQUENCE 482 AA; 52145 MW; F5D493A32F5C7F6C CRC64;

Query Match 21.2%; Score 425.5; DB 2; Length 482;
 Best Local Similarity 29.1%; Pred. No. 1.3e-25;
 Matches 109; Conservative 75; Mismatches 138; Indels 53; Gaps 14

QY 2 TSVAEHLSCGDRPPVLVSSDKGIRHLPKLIQITTSFTFNACGRPIGNGV-TIYN 60
 Db 31 TTDTHKHVKALSPALF-----RAKLWDLQRTWKVTEDKTLACGCRWAGGSVASLW 86
 QY 61 GPKG-SGFGGLRSCGKWCPCPCAGKVGARADEISQVV-----AHQLGTGS-- 106
 Db 87 QAQRAEWGGLQNSHVSVMGSPVASSVICSRRANEVDAAIQAWANGAGLTPAHYRGVSTSP 146
 QY 107 -----VAMVMTMRHTAGQRLHDLMTGLSAWKAATNGRW-- 143
 Db 147 SDRKQGVSNPVVERGISLMTLTLHNSKQSITVEWDAIAGCQAVNTAAWGGARTA 206
 QY 144 TERMYGCDGYRAVEITHGKGNHVVHALLMFSGDVSENILESFSDAMFDRWTSKLV 203
 Db 207 GDRSYGIAHYRAIEVTHGKGNHVVHLLVFLHFDRLVLSVDERSDSLADRVFDRWAAKVR 266
 QY 204 LGFAAPLURNSGGLDVRKIGGEADQV--LAAYLTK-IASGVGMEVSGD-GKSGRHGNRP 259
 Db 267 LGWRASDR-GIDVHVAASSDDAKSIGYTCCKMLSGIAAETTTGQVTRKAGDNRTF 325
 QY 260 WEIADAVAGGDPQALE--LWREFFGMGRRAIAWSRGLRARAGLCAELTDAQIVQ--E 315
 Db 326 FQILGD-LGQYTKRDHALWLEWKGSGRRQTGWSQGTQKVLGIN-ELSDQIDDSLDG 383
 QY 316 ESAPVMVAIIPARSW 330
 Db 384 NDQSEVVVMVGRGAW 398

RESULT 6
 Q9RIQO PRELIMINARY; PRT; 476 AA.
 AC Q9RIQO;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Replication protein.
 GN Name:rep;
 OS Streptomyces natalensis.
 OC Plasmid pSNAL.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=68242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mendes M.V., Aparicio J.F., Martin J.F.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ243257; CAB62261.1; -
 DR GO:0005727; C:extrachromosomal circular DNA; IEA.
 DR GO:0003677; F:DNA binding; IEA.
 DR GO:0006260; P:DNA replication; IEA.
 DR InterPro: IPR000989; Rep.
 DR Pfam: PF01446; Rep_1; 1.
 KW Plasmid.
 SQ SEQUENCE 476 AA; 51890 MW; DEE4C8B3CC2B5B95 CRC64;

Query Match 21.2%; Score 424.5; DB 2; Length 476;
 Best Local Similarity 29.7%; Pred. No. 1.5e-25;
 Matches 121; Conservative 53; Mismatches 148; Indels 85; Gaps 14;

QY 37 ITTSETFNACGRPIGNGVITVNGPKSGF-GGLRSCGKWI CPCGAKGVGAHRADEIS 95
 Db 1 MTTDKALRCRGYALG-GGVTPKVTGGTAYLAGLATCGKVHICPCCGAKIRTSARTVELQ 59

```

QY 96 QV-----VAHOLGTGSGVAMVTMTWHT 117
Db 60 AAGTAWEDIGINGPKIPARAANKOINDRTABDWAIEADGLOAGGLMLTLMRHY 119
QY 118 AGORLHDLTCLSAWKAATN---GRWRTEREMYGCGYVRAVEITHGK-NGWHVHVA 173
Db 120 SRHTLAEILVTOORDAWKALCONAGRWKAKKDYGVVGFVRAVEVYTGANGHWHVH 179
QY 174 LLMFSGDVSENLISFSDAMPDRWTSKLVSLGFAAPLENSG-GLDVRKIGGEADQVLAAY 232
Db 180 LVFEDKPLTPQGDALIEVYEAMSTALQDVGVLPDRHGVRLDLSHGEGGPR--ARY 237
QY 233 LTKIASG-----VGMVEGSGDGKGRHGNRPWELIA-----VDAVGGD-POALELWREFE 281
Db 238 LMKYQDGKAAWTTAAEMTRTDTKAGRDGHRTPFEIARVLLTDAADDDRAQTVRLWQYE 297
QY 282 FSGMGRRAIAISRGIRARAGLGAEL---TDAQIVQEESAPVMVAII-----PA 327
Db 298 TAARGMRLYWSNGRKLALVEALVDTDTDGEIAAEERQGEALAVILADPHWHQIARRKG 357
QY 328 RSWMMIRTCAPYVFEIILGLVEA-GATWENLRDLHLHYELPAADVRPP 373
Db 358 RSLQLLKAERGGQDKVRAVESWGLVW-----GRDVLPP 392

```

RESULT 7

Q74211 ID Q74211 PRELIMINARY; PRT; 525 AA.

AC Q74211; AAS03171; (Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DE Hypothetical protein.

GN OrderedLocusNames=MAP0854;

OS Mycobacterium paratuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium avium complex (MAC).

OX NCBI_TaxID=1770;

RN [1] SEQUENCE FROM N.A.

RC STRAIN=k10;

RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.; Submitted (SEP-2003) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AE017229; AAS03171.1; --

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 525 AA; 56735 MW; DB5E83E68D35FD4 CRC64;

```

Query Match 20.4%; Score 409; DB 2; Length 525;
Best Local Similarity 28.0%; Pred. No. 3e-24;
Matches 125; Conservative 58; Mismatches 136; Indels 128; Gaps 18;

QY 22 DKRGIRHELKPKLQIITSETFNACGR-----PISGVNGVTI-----VNGPKGSGFG 68
Db 62 ERRIRWGARMWLQWASSLKAVRCGRVLHNDVAGDPDG-OGVVKREVDRGMVASLH 120
QY 69 GLRSCGKWCPCACAGKGAHRADEISQV---AHOLGTGSGVAMVTMTWHTAGQRIHDLW 126
Db 121 GLMTGSGWACPRCSAVIANTRAETGATVRECVRG-GRVYLLTLTMRHSRRDGLADLW 179
QY 127 TGLSAWKAATNGRRWTER-----EMYGCDGYVRAVEITHGK-- 164
Db 180 DLSLTAWSRVFGTRNTGTQKERMVQRRRGLALLPEIMGDAERFDIAGVTRVVEATYKPE 239
QY 165 ---NGWHVHVALLMFSGDVSENLIS-----PISGVNGVTI-----FSDAMFRWTSKLVS 203
Db 240 LGHGHWLHIALVFSVTSLSGLLIEGIERTLGRGVNHDWLARNVFAARIHQWWSQGLAK 299
QY 204 LGFAAPLRNSGGLDVRKIGGEADQVLAAYTK-----IASGVGMEVSGS---DGKSGRHG 255
Db 121 GLMTGSGWACPRCSAVIANTRAETGATVRECVRG-GRVYLLTLTMRHSRRDGLADLW 179
QY 127 TGLSAWKAATNGRRWTER-----EMYGCDGYVRAVEITHGK-- 164
Db 180 DLSLTAWSRVFGTRNTGTQKERMVQRRRGLALLPEIMGDAERFDIAGVTRVVEATYKPE 239
QY 165 ---NGWHVHVALLMFSGDVSENLIS-----PISGVNGVTI-----FSDAMFRWTSKLVS 203
Db 240 LGHGHWLHIALVFSVTSLSGLLIEGIERTLGRGVNHDWLARNVFAARIHQWWSQGLAK 299
QY 204 LGFAAPLRNSGGLDVRKIGGEADQVLAAYTK-----IASGVGMEVSGS---DGKSGRHG 255
Db 300 AGCOMP--GSAVADVREIDDEGAEGYVGYLSKATYDVAARIGLEVAGVSTKDARAER-- 355
QY 256 NRAPWEI-----AVDAVG-----GD-----POALELW 277
Db 356 NQTPPEVLANLAESVDARGFGIRTPRHVAVLPAAGNDMAVIDSDTGEVASITAPGQWKV 415
QY 278 REFEGSGMRRRAIAISR-----GLRARAGLGAELTDAQIVQEESAPVMVA 323
Db 416 HEWEQASCGRRQITWSRRRNSPESGREMLNDLLSDRGRSAEASDEEIAVDEVDAB-SVG 474
QY 324 IIPARSWMMIRTCAPYVFEIILGLVEA 350
Db 475 VISRQWYQVFAWRP---GLIVDLLEA 498

```

RESULT 9

Qy	228	VLA-----AYLTIAS-----GVCMEVSGDGKSGRGNRAPWEIAVDVG	268
Db	240	VAAFEAHADVKGWCTASPYLAKQDDSRHWGDAEIAKASTKAGRAGKHFPFALLTLFQD	299
Qy	269	GDQALELWREFFGSMGRRAATWSGLRARAGLGAELTDAQIVEGESAPVVAIIPAR	328
Db	300	GDRKAGHRFLDYAAAMKGRQLSWSAGLAKGVV-TEQSDVEAEQQRDADLLIGHLEIE	358
Qy	329	SNWMITCAPYVEGEILGLVEAGATWENLRDHL	361
Db	359	QWKLTIQAG--LRAKVLDAESNGGWAQI-QHL	388

RESULT 13
Q93L41
ID Q93L41
PRELIMINARY:
PRT: 461 AA.

RC STRAIN-ATCC 15991;
RA Venkova T., Patek M., Nesvera J.;
RT "Characterization of the cryptic plasmid pCC1 from *Corynebacterium*
RT *callunae* and its use for vector construction.";
RL Plasmid 51:54-60(2004).
DR EMBL; AJ308231; CAC59945.1; -.
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006360; P:DNA replication; IEA.
DR InterPro; IPR000989; Rep.
DR Pfam; PF01446; Rep_1; 1.
DR KW SCQUENCE.
DR SC_SEQUENCE 461 AA: 50964 MW: 8747737EC2CE0701 CRC64;

QY	29	ELRPLQOITTTSEFNACGRIPISWN--GVITVNGPKSG-FGGJLRSCKGWICPCFCCAGV	86
Db	66	ERQKAWAITNKKALACGHRWRSGSKGVSLKWANNSSRFAGLONSNSVWASPIPAVAI	125
QY	87	GAHRADEISQVVA--HQL-GTGGSVAMVTMTNRHTTAGORLLDHLDTGLTSLAAKWAANGRRWR	143
Db	126	GKRRVQSAATVAKNWHMDHNDASVLLLTLLTPHAHGTLAKDLSBALKAGNAGIIGTTSWK	185

Db	186	KDREAYRLPWWHKALEITHGQNSFPHPHVLLFCERTILSTTEVEALKARLFDRYAKRLEK	244
Qy	204	LGFAAPLRNNSGGLDVRKVTGGEARQV - LAAYLTK - IAS - -GVGMVEVSGDCKSGRHGNRAP	259

246 DB HGWQRPSEH-GIDLVSQTRDDAIIIMGAYTAKGIAESWNAASEVAGQAFKEAGTNNRP 304
260 QY WEIADVAGGDPQALE-----LWREFFGSMGRATAWSGLRARAGLGAELTDAQIV- 312
305 DB WQILLDDIAAGAGPTPEYRRDVAIWREYEATTRGVKQTSWGTGAKKALSUNV-LKDEEDVAT 363
313 QY -----EQESAPVMVAIIPARSNMMIRTCAPYVFGEILGLVEAGATWENLRDHLHYRLPA 367
364 DB GEVLGDEATEDYVVADIPAKA-----WAELCDDVHKRLUDI 399
368 AD 369 QY

```

Db          400 AN 401
|:
RESULT 14
Q9R881      PRELIMINARY;      PRT;      401 AA.
AC          Q9R881;
DT          01-MAY-2000 (TREMELrel. 13, Created)
DT          01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT          01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE          Replicase.
GN          Name=repB11;
OS          Corynebacterium glutamicum (Brevibacterium flavum).
OG          Plasmid pB11.
OC          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC          Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX          NCBI_TaxID=1718;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 21086;
RX          MEDLINE=96276208; PubMed=8693028;
RA          Ankril S., Reyes O., Leblon G.;
RT          "Electrotransformation of highly DNA-restrictive corynebacteria with
RT          synthetic DNA.";
RL          plasmid 35:62-66(1996).
RN          [2]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 21086;
RA          Favey S., Leblon G., Reyes O.;
RL          Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN          [3]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 21086;
RA          Ankril S., Reyes O., Leblon G.;
RL          Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR          EMBL; AF092037; AAD08690.1; -.
KW          Plasmid.
SQ          SEQUENCE      401 AA; 44481 MW; 5E0590D936132C4D CRC64;

Query Match      16.3%; Score 326; DB 2; Length 401;
Best Local Similarity 29.1%; Pred. No. 1e-17;
Matches 98; Conservative 58; Mismatches 131; Indels 50; Gaps 15;

Qy      34 LQIITSETFNACGR-----PISGVNGVTIVNGPKSGFGLRSCGKWCIFCCAGKV 86
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MYKITNSKALAGCHRRRDEAVAVSWSS-----ASQFGLQNGSHRWGSPLAELV 53

Qy      87 GAHRADEISQVAHQLGTVGSVAVTM-TMRHTAGQRLHDLWTGLSAAKAATNGRRWTE 145
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 MGERIELATKKNHLAAGNALMFVGTVRNRSQSFAQVZAGIKTAYSSMVKTSQWKKE 113

Qy      146 REMYGCQGYVRVETHTG-KNGWHVHVHALLMFSGDVSENILESFSDFAMFDRWTSKLVS 204
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
114 RARYGVETYSDEYVETDSWANGWHLHRNMLLFLDRPLSDDELKAFEDSMFSRWSAGVVK 173

Qy      205 GFAPPLNSG-GLD-VRKIGSEADQVLAAYITKASGVGVGVGVGVGVGVGVGVGVGVGV 262
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
174 GMDAPLREHGVKLDQVSTWGGDAK-NATYLAKE---GMSQELTGSATKTAKSGSYTTFQM 229

Qy      263 AVDAVGDDPOALE-----LWREFFEGSGMRAIAWSRGLRARGAGELTDAQI--- 311
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
230 -LDMLAQSDGDGENDAVLVARWREYEVGSKNLRS-SWSRG--AKRALGDIYIDAVRRE 285

Qy      312 VEQE-----ESAPVMVAIIIPARSWMMIRT 335
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
286 MEELYKLAGEAPERVESTREVAVALYKPDNDKLIQS 322

RESULT 15
REP_STRLI
ID -REP STRLI
AC P22406;
STANDARD;      PRT;      456 AA.

```

```

DT          01-AUG-1991 (Rel. 19, Created)
DT          01-AUG-1991 (Rel. 19, Last sequence update)
DT          05-JUL-2004 (Rel. 44, Last annotation update)
DE          Rep protein.
GN          Name=rep;
OS          Streptomyces lividans.
OG          Plasmid pJ101.
OC          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC          Streptomycineae; Streptomycetaceae; Streptomyces.
OX          NCBI_TaxID=1916;
RN          [1]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=89008081; PubMed=3170481;
RA          Kendall K.J., Cohen S.N.;
RT          "Complete nucleotide sequence of the Streptomyces lividans plasmid
RT          pJ101 and correlation of the sequence with genetic properties.";
RL          J. Bacteriol. 170:4634-4651(1988).
CC          -I- FUNCTION: Essential for the autonomous replication of the plasmid
CC          pJ101.
CC          -----
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC          or send an email to license@isb-sib.ch).
CC          -----
DR          EMBL; M21778; AAA88404.1; -.
DR          PIR; A31844; A31844.
DR          InterPro; IPR000989; Rep.
DR          Pfam; PF01446; Rep_1; 1.
KW          Plasmid; Plasmid partition.
SQ          SEQUENCE      456 AA; 49617 MW; 0376FD9D9429C7DC CRC64;

Query Match      15.8%; Score 316; DB 1; Length 456;
Best Local Similarity 30.5%; Pred. No. 7.5e-17;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

Qy      49 PISGV-----NGVTIVNGPKSGFGLRSCGKWCIFCCAGKVGCAHRADEISQVAHQ 102
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 PASGVIYAQTAAGTSVVL-----GLMRCGRWLCPVCAATIRHKRAEITAAVVEWI 54

Qy      103 GTGSVA-MVTWTRHTAGQRLHDLWTGLSAAKAATNGRR-----MRETE--- 145
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
55 KRGGTAYLVFTARHGHGHTDLADLMDALQGTRKTPDSPRPGAYQRLITGCTWAGRAKD 114

Qy      146 -----REMYGCQGYVRVETHTGK-NGWHVHVHALLMFSGDV-----SENILES 189
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
115 GHRAADREGIRDRIGYVGMIRATEVTVGQINGWHPHIIAIVLGVGRTERGERSAKQIVATF 174

Qy      190 --SDAMFDRWTSKLVSGLGFAAPLR-----NSGGLDVRKIGGEAD-QVLAAYLTAKIA 237
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 EPTGAALDEWQGHWSV-WTAALRKVNPATPDDRHGVDFKLETERDANDLAEYIAKTQ 233

Qy      238 SG--VGMVEVSGGKSGRHNRAPEWI---AVDAVG---DPOA-----IELWREFE 281
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 DGKAPALELARADIKTATGNNVAPFELLGRIGDLTGMDTDDAAGVGSLEWNLRSRWHEYE 293

Qy      282 FGSNGRAIAWSRGLRARGAGELTDA 309
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
294 RATRGRRAIEWTRYLRQMLGLDGGDTEA 321

Search completed: October 23, 2004, 03:57:40
Job time : 112 secs

```


Db	181	GGTCCGAAAGTTCTGGATTGCGAGGCGTTCGTCTCCGCGAAGGCGTGGATCGCCCC	240
Qy	241	TGCTGTGCGGAAAGTCGTGCACATCGTGCACAGCAAAATTTCTCAAGTTGTTCCTCAT	300
Db	241	TGCTGTGCGGAAAGTCGTGCACATCGTGCACAGCAAAATTTCTCAAGTTGTTCCTCAT	300
Qy	301	CAACTCGGACGTGATCTGTTGGCATGCGTGCAGTACGATGACATCGCCCATACAGCTGGTCAG	360
Db	301	CAACTCGGACGTGATCTGTTGGCATGCGTGCAGTACGATGACATCGCCCATACAGCTGGTCAG	360
Qy	361	CGGCTCCACGACCTATGACATGGAATTTTCGCGACGCTGGAAGCTGCACCAACCGTTCGT	420
Db	361	CGGCTCCACGACCTATGACATGGAATTTTCGCGACGCTGGAAGCTGCACCAACCGTTCGT	420
Qy	421	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGACGGATACGTCGCGCTGTTGAAATC	480
Db	421	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGACGGATACGTCGCGCTGTTGAAATC	480
Qy	481	ACTCAGGAAAAACGGCTGCGACGTCCACGTTACCGGCTACTCATGTTTCAGTGGTGAC	540
Db	481	ACTCAGGAAAAACGGCTGCGACGTCCACGTTACCGGCTACTCATGTTTCAGTGGTGAC	540
Qy	541	GTGAGTGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGGACTTCCAAA	600
Db	541	GTGAGTGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGGACTTCCAAA	600
Qy	601	CTCGTATCTCTGGGATTTGCTCGGCCACTACGTAATTCGGGTGCTCGATGTACGAAAG	660
Db	601	CTCGTATCTCTGGGATTTGCTCGGCCACTACGTAATTCGGGTGCTCGATGTACGAAAG	660
Qy	661	ATCGGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCGATCTGGCGTT	720
Db	661	ATCGGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCGATCTGGCGTT	720
Qy	721	GATATGGAGTTGTTAGTGGCGACGAAAGTGGTGCACATGCGCAACCGTGCAACCTGG	780
Db	721	GATATGGAGTTGTTAGTGGCGACGAAAGTGGTGCACATGCGCAACCGTGCAACCTGG	780
Qy	781	GAAATCGCTGTTGATGTCAGTGGCGGGGATCCAAAGCGTTGGAACTGTGGCGGAAATTT	840
Db	781	GAAATCGCTGTTGATGTCAGTGGCGGGGATCCAAAGCGTTGGAACTGTGGCGGAAATTT	840
Qy	841	GAGTTTGGTTCGATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTCGTCGCCGAGCT	900
Db	841	GAGTTTGGTTCGATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTCGTCGCCGAGCT	900
Qy	901	GCTCTTGGGCGAGAACTACAGATGCTCAGATGCTTGGACGAGGAAATCTCCCGCGTTC	960
Db	901	GCTCTTGGGCGAGAACTACAGATGCTCAGATGCTTGGACGAGGAAATCTCCCGCGTTC	960
Qy	961	ATGGTTGGCATCATTCGCGCGGATCGTGGATGATGATTCGGAATTCGTCGCGCTTACGTC	1020
Db	961	ATGGTTGGCATCATTCGCGCGGATCGTGGATGATGATTCGGAATTCGTCGCGCTTACGTC	1020
Qy	1021	TTCCGCGAGATCCTCGGACTCTCGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC	1080
Db	1021	TTCCGCGAGATCCTCGGACTCTCGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC	1080
Qy	1081	TTGCAATATTCGATTTGCCCGCAGCGGATGTCGCGCCCCCGGATAATATTCGTTCCGAGTGA	1140
Db	1081	TTGCAATATTCGATTTGCCCGCAGCGGATGTCGCGCCCCCGGATAATATTCGTTCCGAGTGA	1140
RESULT 2			
AV178757/c			
LOCUS			
Rhodococcus erythropolis cryptic plasmid pAN12, complete sequence.			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacterineae; Nocardiaceae; Rhodococcus.			
1 (bases 1 to 6334)			
Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and Cheng, Q.			
A small cryptic plasmid from Rhodococcus erythropolis: characterization and utility for gene expression			
Appl. Microbiol. Biotechnol. 62 (1), 61-68 (2003)			
22718480			
12835922			
2 (bases 1 to 6334)			
Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and Cheng, Q.			
Direct Submission			
Submitted (11-NOV-2002) CR&D, E.I. Dupont de Nemours Inc., Experimental Station, Wilmington, DE 19880-0328, USA			
Location/Qualifiers			
1. 6334			
/organism="Rhodococcus erythropolis"			
/mol_type="genomic DNA"			
/db_xref="taxon:1833"			
/plasmid="pAN12"			
/note="cryptic plasmid"			
complement(1912..3051)			
/gene="rep"			
complement(1912..3051)			
/gene="rep"			
/note="Rep"			
/codon_start=1			
/transl_table=11			
/product="replication protein"			
/protein_id="AAO31602.1"			
/db_xref="GI:28628257"			
/translation="MTSVSAEHLGSKDPPVLVSKDGKRGIRHELPRKLOITTSFENACGRPIGVNGTIVNGPKSGFGGLRSCGKGNICPCGAKGVGAHRADEISQVVAHQLTGSGVAVTMRHTAGORLHMTGLSAKAAKATNGRWRTERRWYCDGTVRAVEITGKNGHVVHALLMFSGDSENFSDAMFDRWTSKLVSLGFAAPLRNSGGLDVRKIGEDAQVLAAVLTIAISGVMEVSGDGKGRHGNRPWEIADVADGDPQALELWREFPGSMGRRAIWSRGLRAGLGAELTDAQIVEQESAPVMVAITPARSMMWIRTCAPVFGELILGLVEAGATWENLRDLHLHYLRPAADVRPPIISVRK"			
complement(4288..5178)			
/gene="div"			
complement(4288..5178)			
/gene="div"			
/note="Div"			
/codon_start=1			
/transl_table=11			
/product="putative cell division protein"			
/protein_id="AAO31603.1"			
/db_xref="GI:28628258"			
/translation="MDQTDTPITAGWNELAPVPMVDIAKDAAHWLIQKTRSGKSOCTYNLLAAGSNPAVRVGVDPSTVLLAPFVHRRPAEPNIELGNDPDKVLRVLFQVKAESDRIECFWDRRDKISLSPALPLILLVEEPGIIEGADDPDATNGLKPADRYAPRITSILVRQIAQSAKAGIRMLLQARASIVGGNARSFVNMTRLRVDEPESVKMLHFNATPEECALVEGFPVGGGFFDQGLRQMTIRTVRVGEISTVASTVSNADLAIEAALNIDRAQRMTIASEPHLGDIG"			
Query Match			
Best Local Similarity			
Matches 1140; Conservative			
0; Mismatches			
0; Indels			
0; Gaps			
0;			
Qy	1	ATGACACGCGTAAGTGTCTGAACACCTTTCCGCGCAAGACCGGCTCCCGTCTCGTGTGCG	60
Db	3051	ATGACACGCGTAAGTGTCTGAACACCTTTCCGCGCAAGACCGGCTCCCGTCTCGTGTGCG	2992
Qy	61	TCCGATACGCGCGGATCCGGACGACGAACTGCGACCCCAAACTTCAACAATCACCAGTCA	120
Db	2991	TCCGATACGCGCGGATCCGGACGACGAACTGCGACCCCAAACTTCAACAATCACCAGTCA	2932
Qy	121	GAACAATTAAAGCGCTGTGCGCGCGGATTTCTGGCGTGAAACGCTGTGACCATTTGCAAC	180
Db	2931	GAACAATTAAAGCGCTGTGCGCGCGGATTTCTGGCGTGAAACGCTGTGACCATTTGCAAC	2872

```
Qy 181 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 240
Db 2871 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 2812
Qy 241 TGCTGTGCGGAAAAAGTCCGTCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db 2811 TGCTGTGCGGAAAAAGTCCGTCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 2752
Qy 301 CAATCCGGACCTGGATCTGTTGCGATGTTGAGGATGACATGCGGCATACAGCTGGTCA 360
Db 2751 CAATCCGGACCTGGATCTGTTGCGATGTTGAGGATGACATGCGGCATACAGCTGGTCA 2692
Qy 361 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGTTCGACCAACGCTCGT 420
Db 2691 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGTTCGACCAACGCTCGT 2632
Qy 421 GGTGGGCTACGGAACGTTGAAATGTAACGCTGCGACGGATACGTCGCGCTGTTGAAATC 480
Db 2631 GGTGGGCTACGGAACGTTGAAATGTAACGCTGCGACGGATACGTCGCGCTGTTGAAATC 2572
Qy 481 ACTCACGGAAAAACGGCTGGCAGCTCCAGTTCAGCGCTACTCATGTTCAAGTGGTCA 540
Db 2571 ACTCACGGAAAAACGGCTGGCAGCTCCAGTTCAGCGCTACTCATGTTCAAGTGGTCA 2512
Qy 541 GTGAGTGAGAAATCCTCGAATCCTTCTCGGATGCGATGTTTCGATCGGCTGACTTCCAAA 600
Db 2511 GTGAGTGAGAAATCCTCGAATCCTTCTCGGATGCGATGTTTCGATCGGCTGACTTCCAAA 2452
Qy 601 CTCGTATCTCTGGGATTTGTCGCGCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 660
Db 2451 CTCGTATCTCTGGGATTTGTCGCGCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 2392
Qy 661 ATCGCGGCTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTCGATCTGGGCTT 720
Db 2391 ATCGCGGCTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTCGATCTGGGCTT 2332
Qy 721 GGTATGAGGTTGGTAGTGGCGACGGAAGAAAGTGTGACATGCAACCGTGACCCCTGG 780
Db 2331 GGTATGAGGTTGGTAGTGGCGACGGAAGAAAGTGTGACATGCAACCGTGACCCCTGG 2272
Qy 781 GAAATCGCTGTTGATGTCAGTGGGCGGGAATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 840
Db 2271 GAAATCGCTGTTGATGTCAGTGGGCGGGAATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 2212
Qy 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCCGCTGGTCCCGTGGATTCGCTGCGCCGAGCT 900
Db 2211 GAGTTTGGTTCGATGGGACGTCGGGCAATCCGCTGGTCCCGTGGATTCGCTGCGCCGAGCT 2152
Qy 901 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGCGTC 960
Db 2151 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGCGTC 2092
Qy 961 ATGTTTGGGATCAATTCGGGCGGATCGTGGATGATGATTGGACTTGTGCGCTTACGTC 1020
Db 2091 ATGTTTGGGATCAATTCGGGCGGATCGTGGATGATGATTGGACTTGTGCGCTTACGTC 2032
Qy 1021 TTGGGCGAGATCTCGGACCTGTCGAGCTGTCGAGCTGGCGGATTCGGGAAATCTTCGTGATCA 1080
Db 2031 TTGGGCGAGATCTCGGACCTGTCGAGCTGTCGAGCTGGCGGATTCGGGAAATCTTCGTGATCA 1972
Qy 1081 TTGCATTATCATTTGCCGAGCGGATGTCGGGCCCCCGGATAATATATCGGTTCCGCAAGTGA 1140
Db 1971 TTGCATTATCATTTGCCGAGCGGATGTCGGGCCCCCGGATAATATATCGGTTCCGCAAGTGA 1912
```

RESULT 3
AX548646/c
LOCUS AX548646 6334 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 5 from Patent WO02055709.
ACCESSION AX548646
VERSION AX548646.1 GI:25813614
KEYWORDS Rhodococcus erythropolis
SOURCE

ORGANISM Rhodococcus erythropolis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Nocardiaceae; Rhodococcus.
REFERENCE
AUTHORS Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.
TITLE Rhodococcus cloning and expression vectors
JOURNAL Patent: WO 02055709-A 5 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1. .6334
/organism="Rhodococcus erythropolis"
/mol_type="unassigned DNA"
/strain="AN12"
/db_xref="taxon:1833"

ORIGIN

Query Match 100.0%; Score 1140; DB 6; Length 6334;
Best Local Similarity 100.0%; Pred. No. 1e-285;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACACGCGTAAGTGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCTGTCG 60
Db 3051 ATGACACGCGTAAGTGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCTGTCG 2992
Qy 61 TCCGATAGCGCGCATCCGGCAGCAAGCTGCGACCCCAAACTTCAACAAATCACACGTCA 120
Db 2991 TCCGATAGCGCGCATCCGGCAGCAAGCTGCGACCCCAAACTTCAACAAATCACACGTCA 2932
Qy 121 GAAACATTTAAACGCTGTGGCGCGCATTTCTCGCTGACCGGTGACCATTTGTCAAC 180
Db 2931 GAAACATTTAAACGCTGTGGCGCGCATTTCTCGCTGACCGGTGACCATTTGTCAAC 2872
Qy 181 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 240
Db 2871 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 2812
Qy 241 TGCTGTGCGGAAAAAGTCCGTCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db 2811 TGCTGTGCGGAAAAAGTCCGTCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 2752
Qy 301 CAATCCGGACCTGGATCTGTTGCGATGTTGAGGATGACATGCGGCATACAGCTGGTCA 360
Db 2751 CAATCCGGACCTGGATCTGTTGCGATGTTGAGGATGACATGCGGCATACAGCTGGTCA 2692
Qy 361 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGTTCGACCAACGCTCGT 420
Db 2691 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGTTCGACCAACGCTCGT 2632
Qy 421 GGTGGGCTACGGAACGTTGAAATGTAACGCTGCGACGGATACGTCGCGCTGTTGAAATC 480
Db 2631 GGTGGGCTACGGAACGTTGAAATGTAACGCTGCGACGGATACGTCGCGCTGTTGAAATC 2572
Qy 481 ACTCACGGAAAAACGGCTGGCAGCTCCAGTTCAGCGCTACTCATGTTCAAGTGGTCA 540
Db 2571 ACTCACGGAAAAACGGCTGGCAGCTCCAGTTCAGCGCTACTCATGTTCAAGTGGTCA 2512
Qy 541 GTGAGTGAGAAATCCTCGAATCCTTCTCGGATGCGATGTTTCGATCGGCTGACTTCCAAA 600
Db 2511 GTGAGTGAGAAATCCTCGAATCCTTCTCGGATGCGATGTTTCGATCGGCTGACTTCCAAA 2452
Qy 601 CTGATATCTCTGGGATTTGTCGCGCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 660
Db 2451 CTGATATCTCTGGGATTTGTCGCGCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 2392
Qy 661 ATCGCGGCTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTCGATCTGGGCTT 720
Db 2391 ATCGCGGCTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTCGATCTGGGCTT 2332
Qy 721 GGTATGAGGTTGGTAGTGGCGACGGAAGAAAGTGTGACATGCAACCGTGACCCCTGG 780
Db 2331 GGTATGAGGTTGGTAGTGGCGACGGAAGAAAGTGTGACATGCAACCGTGACCCCTGG 2272
Qy 781 GAAATCGCTGTTGATGTCAGTGGGCGGGAATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 840


```
Db 2271 GAAATCGCTGTTGATGAGTGGGCGGGATCCACAAGGTTGGAACTGTGCGGAGATTT 2212
Qy 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGATTCGGTGGCCGAGCT 900
Db 2211 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGATTCGGTGGCCGAGCT 2152
Qy 901 GGTCTTGGGGGAGAACTAACAGATGCTCAGATGCTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 2151 GGTCTTGGGGGAGAACTAACAGATGCTCAGATGCTTGAGCAGGAGAAATCTGCCCGGTC 2092
Qy 961 ATGGTTGCGATCATTTCCGGCGCGATCGTGGAATGATGATTCGGACTTTGCGCCCTTACGTC 1020
Db 2091 ATGGTTGCGATCATTTCCGGCGCGATCGTGGAATGATGATTCGGACTTTGCGCCCTTACGTC 2032
Qy 1021 TTCGGCGAGATTCCTCGGACTCGTCGAAGCTCGCGGCACTTCGGGAAATCTTCGTGATCAC 1080
Db 2031 TTCGGCGAGATTCCTCGGACTCGTCGAAGCTCGCGGCACTTCGGGAAATCTTCGTGATCAC 1972
Qy 1081 TTGCATTATCGATTGCCCGGAGCGGATGTGGGCCCCCGAATAATATCGGTTTCGCAAGTGA 1140
Db 1971 TTGCATTATCGATTGCCCGGAGCGGATGTGGGCCCCCGAATAATATCGGTTTCGCAAGTGA 1912

RESULT 4
AX548648/c 9652 bp DNA linear PAT 27-NOV-2002
LOCUS AX548648
DEFINITION Sequence 7 from Patent WO02055709.
ACCESSION AX548648
VERSION AX548648.1 GI:25813616
KEYWORDS
SOURCE Shuttle vector pRHBR17
ORGANISM Shuttle vector pRHBR17
artificial sequences; vectors.
REFERENCE
1 Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.
Rhodococcus cloning and expression vectors
Patent: WO 02055709-A 7 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
Location/Qualifiers
source
/organism="Shuttle vector pRHBR17"
/mol_type="unassigned DNA"
/db_xref="taxon:214938"

ORIGIN
Query Match 100.0%; Score 1140; DB 6; Length 9652;
Best Local Similarity 100.0%; Pred. No. 1e-285;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACAGGTTAAGTGTGACACCTTTCCGGCAAGACCGGCTCCCGTCCCTCGTGTGCG 60
Db 5052 ATGACAGGTTAAGTGTGACACCTTTCCGGCAAGACCGGCTCCCGTCCCTCGTGTGCG 4993

Qy 61 TCCGATAAGCGGGATCCGGACGAACTGCGACCAACAACTTCAACAAATACCAACGTCA 120
Db 4992 TCCGATAAGCGGGATCCGGACGAACTGCGACCAACAACTTCAACAAATACCAACGTCA 4933

Qy 121 GAAACATTAAACGCTGTGGCGGCGGATTTCTGCGTGAACGGTGTGACCAATTGTCAAC 180
Db 4932 GAAACATTAAACGCTGTGGCGGCGGATTTCTGCGTGAACGGTGTGACCAATTGTCAAC 4873

Qy 181 GGTCCGAAGTTCGGATTCCGAGCGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 240
Db 4872 GGTCCGAAGTTCGGATTCCGAGCGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 4813

Qy 241 TGCTGTGGGGGAAAAAGTCGGTGACATCGTGAGAGCAAAATTTCTCAAGTTCTTGCTCAT 300
Db 4812 TGCTGTGGGGGAAAAAGTCGGTGACATCGTGAGAGCAAAATTTCTCAAGTTCTTGCTCAT 4753

Qy 301 CAACTCGGAGCTGGATCTGTTGGATGGTGAAGATGACCATGGCCATACAGCTGCTGAG 360
Db 4752 CAACTCGGAGCTGGATCTGTTGGATGGTGAAGATGACCATGGCCATACAGCTGCTGAG 4693
```

```
Qy 361 CGGCTCCACGACCTATGAACTGGAACTTTCCGACGCTTGGAAAGCTGGACCAACCGTGGT 420
Db 4692 CGGCTCCACGACCTATGAACTGGAACTTTCCGACGCTTGGAAAGCTGGACCAACCGTGGT 4633
Qy 421 CGTTGGGCTAGCGAACTGAAATGTAACGGTTCGACGAGTACGTCGCGCTGTTGAAATC 480
Db 4632 CGTTGGGCTAGCGAACTGAAATGTAACGGTTCGACGAGTACGTCGCGCTGTTGAAATC 4573
Qy 481 ACTCACGGAAAAAAGCGGCTGGCAGCTGCAAGTTACGCGCTACTCATGTTCAAGTGGTGC 540
Db 4572 ACTCACGGAAAAAAGCGGCTGGCAGCTGCAAGTTACGCGCTACTCATGTTCAAGTGGTGC 4513
Qy 541 GTGAGTGGAGAACATCTTCGAAATCTCTCGATCGGATGTTCCGATCGGTGACTTCCAAA 600
Db 4512 GTGAGTGGAGAACATCTTCGAAATCTCTCGATCGGATGTTCCGATCGGTGACTTCCAAA 4453
Qy 601 CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGGTGGTCTCGATGTGAGAAAG 660
Db 4452 CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGGTGGTCTCGATGTGAGAAAG 4393
Qy 661 ATCGGCGGTGAGCTGATCAAGTTCTCGCTCGGATCTGACGAAAAATTCGATCTGGCGCT 720
Db 4392 ATCGGCGGTGAGCTGATCAAGTTCTCGCTCGGATCTGACGAAAAATTCGATCTGGCGCT 4333
Qy 721 GGTATGAGAGTTCGTAGTGGCGACGAAAAAGTGGTGCACATGGCAACCGTGCAACCTGG 780
Db 4332 GGTATGAGAGTTCGTAGTGGCGACGAAAAAGTGGTGCACATGGCAACCGTGCAACCTGG 4273
Qy 781 GAAATCGCTGTTGATGACAGTGGGCGGGATCCAAAGCTTTGGAATCTGCGGAGAAATTT 840
Db 4272 GAAATCGCTGTTGATGACAGTGGGCGGGATCCAAAGCTTTGGAATCTGCGGAGAAATTT 4213
Qy 841 GAGTTTGGTTCGATGGGACGTCGGCAATCGGCTGGTCCCGTGGATTCGCTGCCGAGCT 900
Db 4212 GAGTTTGGTTCGATGGGACGTCGGCAATCGGCTGGTCCCGTGGATTCGCTGCCGAGCT 4153
Qy 901 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 4152 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 4093
Qy 961 ATGGTTGCGATCATTTCCGGCGCGATCGTGGAATGATGATTCGAGCTTTGGCGCTTACGTC 1020
Db 4092 ATGGTTGCGATCATTTCCGGCGCGATCGTGGAATGATGATTCGAGCTTTGGCGCTTACGTC 4033
Qy 1021 TTCGGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTGGAAAAATCTTCGATGATCAC 1080
Db 4032 TTCGGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTGGAAAAATCTTCGATGATCAC 3973
Qy 1081 TTGCATTATCGATTGCCCGGAGCGGATGTGGGCCCCCGAATAATATCGGTTTCGCAAGTGA 1140
Db 3972 TTGCATTATCGATTGCCCGGAGCGGATGTGGGCCCCCGAATAATATCGGTTTCGCAAGTGA 3913

RESULT 5
AX548647/c 11241 bp DNA linear PAT 27-NOV-2002
LOCUS AX548647
DEFINITION Sequence 6 from Patent WO02055709.
ACCESSION AX548647
VERSION AX548647.1 GI:25813615
KEYWORDS
SOURCE Shuttle vector pRHBR17
ORGANISM Shuttle vector pRHBR17
artificial sequences; vectors.
REFERENCE
1 Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.
Rhodococcus cloning and expression vectors
Patent: WO 02055709-A 6 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
Location/Qualifiers
source
/organism="Shuttle vector pRHBR17"
/mol_type="unassigned DNA"
```

Db	5681	ATGGTTGGCATCATTTCCGGCGCGATCGTGGATGATGATTCGGAACTTGTGCGCTTACGTC	5682
Qy	1021	TTCCGGCGAGATCCTCGGACTCGTCGAAGCTGCGCGACTTGGGAAATCTTTCGGTATCAC	1081
Db	5621	TTCCGGCGAGATCCTCGGACTCGTCGAAGCTGCGCGACTTGGGAAATCTTTCGGTATCAC	5562
Qy	1081	TTGCATTATTCGATTCCCGCAGCGGATGTGCGGCCCGCGATAATATCGTTTCGCAAGTGA	1140
Db	5561	TTGCATTATTCGATTCCCGCAGCGGATGTGCGGCCCGCGATAATATCGTTTCGCAAGTGA	5502
RESULT 6			
AY180162/c	AY180162	Shuttle vector pRHBR17, complete sequence.	circular SYN 23-JUL-2003
LOCUS	AY180162	Shuttle vector pRHBR17, complete sequence.	
DEFINITION	AY180162	Shuttle vector pRHBR17	
ACCESSION	AY180162	Shuttle vector pRHBR17	
VERSION	AY180162.1	GI:30313707	
KEYWORDS		artificial sequences: vectors.	
SOURCE		1 (bases 1 to 11241)	
ORGANISM		Kostichka,K., Tao,L., Bramucci,M., Tomb,J.-F., Nagarajan,V. and	
REFERENCE		Cheng,Q.	
AUTHORS		A small cryptic plasmid from Rhodococcus erythropolis:	
TITLE		characterization and utility for gene expression	
JOURNAL		Appl. Microbiol. Biotechnol. 62 (1), 61-68 (2003)	
MEDLINE		22718480	
PUBMED		12835922	
REFERENCE		2 (bases 1 to 11241)	
AUTHORS		Kostichka,K., Tao,L., Bramucci,M., Tomb,J.-F., Nagarajan,V. and	
TITLE		Cheng,Q.	
JOURNAL		Direct Submission	
Submitted (14-NOV-2002)		CR@D, E. I. Dupont de Nemours Inc.,	
Experimental Station, Wilmington, DE 19880-0328, USA			
FEATURES			
source			
1..11241			
Location/Qualifiers			
/organism="Rhodococcus erythropolis"			
/mol_type="genomic DNA"			
/db_xref="taxon:214938"			
/focus			
/note="Rhodococcus - Escherichia coli shuttle vector; Ssp			
I linearize pAN12 cloned into the PvuII site of pBR328"			
Join(9925..11241,1..3590)			
/organism="Cloning vector pBR328"			
/mol_type="genomic DNA"			
/db_xref="taxon:221946"			
3591..9924			
/organism="Rhodococcus erythropolis"			
/mol_type="genomic DNA"			
/db_xref="taxon:1833"			
/plasmid="pAN12"			
/note="cryptic plasmid"			
ORIGIN			
Query Match			
Best Local Similarity 100.0%; Score 1140; DB 12; Length 11241;			
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGACGAGGTAAAGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG	60
Db	6641	ATGACGAGGTAAAGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG	6582
Qy	61	TCGATAAGCGGGCATCCGCGACGACCTGCGACCCAACTTCAACAAATCACACGTCA	120
Db	6581	TCGATAAGCGGGCATCCGCGACGACCTGCGACCCAACTTCAACAAATCACACGTCA	6522
Qy	121	GAACATTTAAACGCTGTGGCGCGCGATTTCTGCGGTGAACGGTGTGACCATTTGTCAAC	180
Db	6521	GAACATTTAAACGCTGTGGCGCGCGATTTCTGCGGTGAACGGTGTGACCATTTGTCAAC	6462
Qy	181	GGTCCGAAGGTTCTGGATTCCGAGGCTTGTCTTCGCGAAAGGGCTGGATCTGCCCC	240
Db	6461	GGTCCGAAGGTTCTGGATTCCGAGGCTTGTCTTCGCGAAAGGGCTGGATCTGCCCC	6402
Qy	241	TGCTGTGCGGAAAGTGGTGCACATCGTCAGACGAAATTTCTCAAGTTGTTCAT	300
Db	6401	TGCTGTGCGGAAAGTGGTGCACATCGTCAGACGAAATTTCTCAAGTTGTTCAT	6342
Qy	301	CAACTCGGACTGGATCTGTTGCGATGTTGACGATGACCATCGGCCATACAGCTGGTCAG	360
Db	6341	CAACTCGGACTGGATCTGTTGCGATGTTGACGATGACCATCGGCCATACAGCTGGTCAG	6282
Qy	361	CGGCTCCAGACTATGACGTGGATCTTTCGGAGCTGGAAAGTGGCAACAAAGTGGT	420
Db	6281	CGGCTCCAGACTATGACGTGGATCTTTCGGAGCTGGAAAGTGGCAACAAAGTGGT	6222
Qy	421	CGTTGGCGTACGGAACGTGAAATGTACGGCTGCGACGATACGTCGGCGTGTGAAATC	480
Db	6221	CGTTGGCGTACGGAACGTGAAATGTACGGCTGCGACGATACGTCGGCGTGTGAAATC	6162
Qy	481	ACTCACGGAAAAACCGCTGGCGACCTCCACCGTTCACGGCTPACTCATGTTCAAGTGTGAC	540
Db	6161	ACTCACGGAAAAACCGCTGGCGACCTCCACCGTTCACGGCTPACTCATGTTCAAGTGTGAC	6102
Qy	541	GTGAGTGAGAACATCCTGAAATCTTCGAGATGCGATGTTGATGCGGTGACATTCGAA	600
Db	6101	GTGAGTGAGAACATCCTGAAATCTTCGAGATGCGATGTTGATGCGGTGACATTCGAA	6042
Qy	601	CTCGTATCTCGGATTTGCTGCGCACTAGCTAATTCGGGTGCTTCGATGTACGAAAG	660
Db	6041	CTCGTATCTCGGATTTGCTGCGCACTAGCTAATTCGGGTGCTTCGATGTACGAAAG	5982
Qy	661	ATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTTGATCTGGCGTT	720
Db	5981	ATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTTGATCTGGCGTT	5922
Qy	721	GATATGAGGTGTTGATGTCGCAACGAAAAAGTGGTGCACATGCGCAACCGTGCACCGTGG	780
Db	5921	GATATGAGGTGTTGATGTCGCAACGAAAAAGTGGTGCACATGCGCAACCGTGCACCGTGG	5862
Qy	7		

Db 6461 GGTCCGAAAGTTCTGATTCGGAGGCTTGGTTCCTCGCGAAAGGCTGATCTGCCCC 6402
Qy 241 TGCTGTGCGGAAAGTCGGTGCACTGTGCGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db 6401 TGCTGTGCGGAAAGTCGGTGCACTGTGCGACGAAATTTCTCAAGTTGTTGCTCAT 6342
Qy 301 CAATCGGAGTCGATCTGTTGCGATGGTGAGGATGACCATGCGCCATACAGCTGTCAG 360
Db 6341 CAATCGGAGTCGATCTGTTGCGATGGTGAGGATGACCATGCGCCATACAGCTGTCAG 6282
Qy 361 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCCTCGAAAGCTGCGACCAACGGTCGT 420
Db 6281 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCCTCGAAAGCTGCGACCAACGGTCGT 6222
Qy 421 CGTTGGCGTACGGAACGTTGAATGTACGGCTGCGACGGATACGTGCGCGCTGTTGAATC 480
Db 6221 CGTTGGCGTACGGAACGTTGAATGTACGGCTGCGACGGATACGTGCGCGCTGTTGAATC 6162
Qy 481 ACTCACGGAAGAAACGGCTGCGACGCTCCACGCTTACGCGCTACTCATGTTTCAGTGGTGAC 540
Db 6161 ACTCACGGAAGAAACGGCTGCGACGCTCCACGCTTACGCGCTACTCATGTTTCAGTGGTGAC 6102
Qy 541 GTGAGTGAGAACATCTCGAATCTCTTCGGATCGATGCGATGTTTCGATCGGTGGACTTCCAAA 600
Db 6101 GTGAGTGAGAACATCTCGAATCTCTTCGGATCGATGCGATGTTTCGATCGGTGGACTTCCAAA 6042
Qy 601 CTCGTATCTCTGGGATTTGTCGCGCACTACGTAATTCGGGTGCTCGATGTCAGAAAG 660
Db 6041 CTCGTATCTCTGGGATTTGTCGCGCACTACGTAATTCGGGTGCTCGATGTCAGAAAG 5982
Qy 661 ATCCGCGTGAGCTGATCAAGTTCCTCGCTCGGTATCTGACGAAATTCGATCTGGCGTT 720
Db 5981 ATCCGCGTGAGCTGATCAAGTTCCTCGCTCGGTATCTGACGAAATTCGATCTGGCGTT 5922
Qy 721 GGTATGGAGGTTGTAGTGGCGACGGAAGAGTGTGACATGCGCAACCGTGCACCCCTGG 780
Db 5921 GGTATGGAGGTTGTAGTGGCGACGGAAGAGTGTGACATGCGCAACCGTGCACCCCTGG 5862
Qy 781 GAAATCGCTGTGTGATGCGATGCGGCGGATCCAAAGCGTTGGAACTGTGGCGGAAATTT 840
Db 5861 GAAATCGCTGTGTGATGCGATGCGGCGGATCCAAAGCGTTGGAACTGTGGCGGAAATTT 5802
Qy 841 GAGTTTGGTTCGATCGGAGCTCGGCAATCGCGTGTCCCGTGGATTCGTCGCCGAGCT 900
Db 5801 GAGTTTGGTTCGATCGGAGCTCGGCAATCGCGTGTCCCGTGGATTCGTCGCCGAGCT 5742
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 5682
Qy 961 ATGTTTGGATCATTTCCGCGCGGATCGTGATGATGATTCGGAATTCGCGCTTACGTC 1020
Db 5681 ATGTTTGGATCATTTCCGCGCGGATCGTGATGATGATTCGGAATTCGCGCTTACGTC 5622
Qy 1021 TTCGCGAGATCCTCGACTCTCGAAGCTGCGCGGACTTGGGAAATCTTCGTGATCAC 1080
Db 5621 TTCGCGAGATCCTCGACTCTCGAAGCTGCGCGGACTTGGGAAATCTTCGTGATCAC 5562
Qy 1081 TTGCATTTATCGATTGCCCGCAGCGGATGTCGGGCCCCCGGATATTCGTTTCGCAAGTGA 1140
Db 5561 TTGCATTTATCGATTGCCCGCAGCGGATGTCGGGCCCCCGGATATTCGTTTCGCAAGTGA 5502

RESULT 7

AY150274 3540 bp DNA circular BCT 02-JAN-2003
LOCUS Propionibacterium granulosum cryptic plasmid pPG01, complete
DEFINITION sequence.

ACCESSION AY150274
VERSION AY150274.1 GI:27465054

KEYWORDS Propionibacterium granulosum
SOURCE Propionibacterium granulosum
ORGANISM

ORIGIN

Query Match 8.3%; Score 94.4; DB 1; Length 3540;
Best Local Similarity 56.2%; Pred. No. 2.1e-13;
Matches 198; Conservative 0; Mismatches 151; Indels 3; Gaps 1;
Qy 181 GGTCCGAAAGTTCTGGAATTCGGAGCCCTTCGTTCTCGGAAAGGCTGATCTGCCCC 240
Db 2267 GGTCTATGTTGGTGGCGGTTATGCCGCTTGGCGACGTGCGGAAGTGTGTGGGCTTGCCT 2326

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Propionibacterineae; Propionibacteriaceae; Propionibacterium.
1 (bases 1 to 3540)
Farrar,M.D. and Holland,K.T.
Isolation and characterisation of a cryptic plasmid from the human
skin commensal Propionibacterium granulosum
Unpublished
2 (bases 1 to 3540)
Farrar,M.D. and Holland,K.T.
Direct Submission
Submitted (13-SEP-2002) Skin Research Centre, Division of
Microbiology, University of Leeds, Leeds LS2 9JT, UK
FEATURES
Location/Qualifiers
1..3540
/organism="Propionibacterium granulosum"
/mol_type="genomic DNA"
/strain="PF283"
/db_xref="taxon:33011"
/plasmid="cryptic plasmid pPG01"
74..280
/codon_start=1
/transl_table=11
/product="Orf3"
/protein_id="AAN78121.1"
/db_xref="GI:27465055"
/translation="MGQYGLNRGFYNGSHNEIIGGYPMNGRVIVWPDALLSAMPL
SSGFSGHVVEPRAEVFHRCADHD"
449..1792
/gene="tra"
449..1792
/gene="tra"
/note="plasmid transfer protein"
/codon_start=1
/transl_table=11
/product="Tra"
/protein_id="AAN78122.1"
/db_xref="GI:27465056"
/translation="WAGERTVLTGVYWLRLPVLILLGVAFACWCFVSHLVYVIGFI
VAALAVAFPPSVQVQPKMVPARYVYVWMSLAKAPKPIESYGGDRIVYRPGIOWI
RSDRHVLHVLVRPAGLADSAYLEKGAETQQRKKSMTCLIVPAEAGLDLIPR
DATAGDLPVAPTSWNVPVGVKPDGSEVVDLSPHSHILVSKTRSGKSSFVYGLL
DOMRHLPTVAGVDPTGILFNLGDMGWDALRSKRITNDADAADVQVLSMTDEM
RRIYLLCEHDKWNRNDFSDPQRLLIIVILEYPGILERLQNFDSARGARSSDRFA
SKAALGNRIAEKAGVGVLLLVTRPDAXIIGPLRAQLTTRVTPAQDSDGLRMSH
PELSBOVKQNSWASGVGFIEADGVILTRFSYRAELTDLHRPGASVGQIDLIQ"
1979..3280
/gene="rep"
1979..3280
/gene="rep"
/note="plasmid replication protein"
/codon_start=1
/transl_table=11
/product="Rep"
/protein_id="AAN78123.1"
/db_xref="GI:27465057"
/translation="WLNITVLSKAQPHRLCSEHPCAARSRORAAAGDRIGTTRAT
FPVSDSPCKRNEGRHRYEMRDLGNPQVMPLEVRKCGAVPVQSRIALMAGHGAG
YLAGATCSGVWACPVCAKISAHRRDELARVVQVAVGLGFKVMSLTLTQRHAGQDLA
ELWASLGSWNAVTSRRRWQEFCAQLGVQVWKAVEVTHGSHGHVHVHLVLSKQDP
TSVDTKIRHRKQRRRTPTYPEVQRPEDFAERWSRGLRKRGVDFIAGSGGLDWTA
DSGDEALGRYVAKWNSVDGLANEATLGKFKARNRPTFQILEDFTDTSGLDLR
LWRTVVSASHRKALTWKSLRDWAGMESEDEQVAAQDQCGEAVLFDHDAWQIR
TAGRAFLDLELHGSBGVYAWLKKRRIHYEILPVPWSTST"


```
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingorsoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE 21583737
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingorsoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source
1..125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S20 and WI-12410"
/notes="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement(35581..35746)
repeat_region
/rpt_type=tandem
/rpt_unit="ctg"
complement(<36507..>36887)
/gene="JPH3"
/feature="JPH3"
/notes="synonym: JP3"
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"
/notes="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/db_xref="GI:17646244"
/translation="MSSGGRFDDGSSYCGWEDGKARHGVCCTGPKGGEYTGSSWS
HGFEVLGYVTPSGNTYGTWAGKRGHIGLESKQWYKGEWTHGFKRGYVRECA
NGAKYEGTWSNGLQDGYGTETYSYG"
ORIGIN
Query Match 4.3%; Score 49; DB 9; Length 125020;
Best Local Similarity 12.3%; Pred. No. 0.13;
Matches 114; Conservative 385; Mismatches 419; Indels 10; Gaps 4;
Qy 147 GATTTCGCGTGAACGGGTGACCATTTGTCACCGTTCGGAAGTCTTCGATTCGGAGG 206
Db 17869 GATCTCMKAGKWCMTGKSKGRKYASKGRKKGTTTWRKMSAWSRKMRWYMSMR 17810
Qy 207 CCTTCGTTCTCGGAAAGGGTGGATCTGCCCTCTGTCGGGAAAGTCCGTCACA 266
Db 17809 CAGMGAACGASRSKMKCWGGRKRSAGSKSRGGTGYNNKKGKSGMSKWKGSSTS 17750
Qy 267 TCGTCAGACGAAATTTCTCAAGTTTGTTCATCACTCGGATCGGATCTGTTGCGAT 326
Db 17749 RRRGSSAKSCSYW--GMSMCMSCMSAKSYWMCYCYMYRMSYMYYYKYCYSCMGMSST 17692
Qy 327 GGTGACGATGACATCGGCCATACAGCTGTCAGCGGTCCACGACCTATGACATGGACT 386
Db 17691 SYSYCCMKMSWGSYCKMKYSGWSSYSTSMGYSTSTCKKYKCYSCSSMYKCKTSKSY 17632
Qy 387 TTCGGCAGCCTGGAAAGCTCGCACCAACAGTCTGCTGCTGGCGTACGGAACGTGAAATGTA 446
Db 17631 RRKESYYWGGGRKAKKYYCAGRRRMSYWKCAKWMYSYCCWMSYCWYYSKTYKS 17572
Qy 447 CGGCTGCGACGATACGTGCGCGTGTGAAATCACTACGGAAGAAACGGCTGCGACGT 506
Db 17571 STCYKRGYWGSKTYCTYAGKSRSMYYCMMRSSKSSWSMSMARSSWCMGAGYRRSK 17512
Qy 507 CCAGTTTACCGCGTACTCATGTTCACTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 566
Db 17511 RSAGWAGMRSSKGRSTGMRKACSKTKTSGTSGTSGTSGTSGTSGTSGTSGTSGTSGT 17457
Qy 567 CTCGGATCGCATGTTTCGATCGTGGACTTCCAAACTCGTATCTCTCTGGATTTGCTGGCC 626
Db 17456 KTCYCMKYKYRKTSMCWYMYKSGYKRYKRCMKKKGCTGYRGMSSKSGYKSMRGM 17397
Qy 627 ACTACGTAATTCGGGTGCTCTCGATGATACGAAAGATCGCGGTGAAGTGTATCAAGTTCT 686
Db 17396 RGSYSTSCWKSVCWYMKCMKTYMYKRRSRMRGSMSSKGRGAGYAGRGYCYSSMSTR 17337
Qy 687 CGCTGCGTATCTGACGAAATTCATCTGGCGTTGTTGATGAGGTGTGATGCGCAGCG 746
Db 17336 KRRSKCYSYKKGKMGKMGWGMKGRGSKTYWSSMKMKKRSWKSCKYKSYKSGRRSK 17277
Qy 747 AAAAAGTGTGTCACATCGCAACCGTCCCTCGGAAATCGCTTGTGATGAGTGGCGG 806
Db 17276 GWGSRSTSKAKSSMRMAGSKCTYGSYNNRRNNRNGKTCNTYMYRRSRAMNGNAAAA 17217
Qy 807 GGATCCACAGCGTTGGAATCTGTGGCG--AGAATTGAGTTGTTGTTGATGCGGACGTGCG 864
Db 17216 GCTTCCCATNTGGGGGAAAAAGGCGSASRASCYKGRMSKSKYCSYSGTRRRCKSKCR 17157
Qy 865 GCAATCGGTGTCCTCGGTGATTCGTCGCCGAGTGTCTTGGGCGAAGTCAACAGAT 924
Db 17156 GSRGKMGMTGRSGGKTSYSAKSGRGCTTCWNGRKGKRGKMSRMMK-MYKSYRRKR 17098
Qy 925 GCTCAGATCGTTGAGCAGGAAGAATCTGCCCGGTTCATGTTGCGATCATTCGCGCGCA 984
Db 17097 WMTCKMCMYMYMAYCRSMCMCKSCCGCYSMGMSYSGYSGYKSGMSYMWRSYVS 17038
Qy 985 TCGTGGATGATGATTCGACTTTGTCGCTTACGCTTCGCGGAGATCTCGGATCTGCTC 1044
Db 17037 KRSTSKAWRSSKRGMTGTGRTYKGGGSGYKGGGSGYKGGGSGGKSGKTRSSAGKSSY 16978
Qy 1045 GAAGTCGCGCGACTTCGGGAAATCTTC 1072
Db 16977 KMSCCARYKSNYSKYCSYRMMWCMYY 16950
RESULT 13
PER6662/c 2051 bp DNA circular BCT 08-JUN-1998
LOCUS Propionibacterium freudenreichii plasmid pLME108 rep gene.
DEFINITION AJ006662
ACCESSION AJ006662.1 GI:3212127
VERSION AJ006662.1
KEYWORDS putative; rep gene.
SOURCE Propionibacterium freudenreichii
ORGANISM Propionibacterium freudenreichii
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Propionibacteriaceae; Propionibacterium.
REFERENCE 1
AUTHORS Dasen,G.H., Miescher,S., Teuber,M. and Meile,L.
TITLE Molecular analysis of plasmid pLME108, a plasmid isolated from
Propionibacterium freudenreichii
JOURNAL Unpublished
AUTHORS Dasen,G.H.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-1998) Dasen G.H., Institute of Food Science,
Laboratory of Food Microbiology, ETH Zurich, LFO G24.2,
Schmelzbergstr. 9, CH-8092 Zurich, SWITZERLAND
FEATURES
source
1..2051
/organism="Propionibacterium freudenreichii"
```



```
/mol_type="genomic DNA"
/db_xref="taxon:1744"
/lab_host="DPF2"
/plasmid="pLME108"
complement (223. .1434)
/gene="rep"
complement (223. .1434)
/gene="rep"
/codon_start=1
/transl_table=11
/product="putative rep protein"
/protein_id="CAA07175.1"
/db_xref="GI:3212128"
/db_xref="GOA:O69764"
/db_xref="TREMBL:O69764"
/translacion="MVVDARARSGVRCGGGSPAPLGLDHNANNVSPSAVPA
GEVADKRRFESRYLWRLHTSLKRVAFGRVAASAVSVGRSCDRAGFAGLQSC
GVNACVPCNAKIMARGLGELGAVEWTXKHGGRVAPMTFTVRSRDXSLTAVWDGVA
SGNRVTSKGKWTSDQLRHGVGVFVRVVEVTHGRNGWHVHLVFLVGVDFGDAALAH
RSMFRWEGVLAAGLCTPLAARDVOQMSNAATGLDLHARVLSKAQFQGIKIGHELTHS
QKTSARSLSTRTWELVDGAAGLAKEVLEWFEWEGSGRGRQIGWSAGLDRKLGUM
VESDDVIAAEVGSVADTVALITGDGWRRLVGQOKLYECLRSCELGQGLARWQJE
HGIEHELVEVD"

ORIGIN
Query Match 4.1%; Score 47.2; DB 1; Length 2051;
Best Local Similarity 49.4%; Pred. No. 0.42; 178; Indels 3; Gaps 2;
Matches 177; Conservative 0; Mismatches 178; Indels 3; Gaps 2;

Qy 194 CTGATTTCGGAGGCTTCGTTCTGCGGAAAGGCTGGATCGCCCTGCTGTCGGCGAA 253
Db 1154 CCGGTTTCGTTGCTTCAGAGCTGCGATCGGCTCGCCGCTGCTGTAACCGGA 1095

Qy 254 AAGTCGTGCACATCGTCGCGAGACGAATTTCTCAAGTTGTG-CTCATCACTCGGAC- 311
Db 1094 AGATCATGGCTCGCGTGGCCCTTGAACCTCGCGCGCGGCTCGAGACCTGGACTAAGCACG 1035

Qy 312 -TGCATCTGTTGCCATGTGACAGATGACCATGCCATACAGCTGGTCAGCGGCTCCACG 370
Db 1034 GTGTCGTGTGGCTTCATCAGCTTCAGGTTCCGGTTCGGCACTCTCGTAAGATTCCTTGACGG 975

Qy 371 ACCTATGACCTGGACTTTCGGCAGCCTCGGAAGCTGCGACCAACGGTCGTGTGGCGTA 430
Db 974 CTGTTGGATGTGTGCTTCGGGGTGGCGGAGTCACTTCGGGCAAGGCTGGAGCT 915

Qy 431 CGGAACGTGAATGATCGGCTGCGACGAGTACGTGCGCGCTGTGGAATCACTCACCGAA 490
Db 914 CTGACCAAGCTCCGTCACGGGGTCGAGGCTTCGTGCGTGTGTCGAGTTACGCACGGC 855

Qy 491 AAAACGCTGGCAGCTCCACGTTTCAACGCGCTACTCATGTTCAAGTGGTGAGTGATGA 548
Db 854 GCAATGGTGCGCATGTGCACTGTCATGTGCTCGTGTCTCGGTCCGGGACTTCGGCGA 797

RESULT 14
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
21583737
```

```
11694876
2 (bases 1 to 125020)
Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
Location/Qualifiers
1. .125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/notes="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement (35581. .35746)
/rpt_type=tandem
/rpt_unit="ctg"
complement (<36507. .>36887)
/gene="JPH3"
/notes="synonym: JP3"
complement (<36507. .>36887)
/gene="JPH3"
product="junctophilin 3"
complement (<36507. .36887)
/gene="JPH3"
/notes="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AA140941.1"
/db_xref="GI:17646245"
/translacion="MSGGRFNFDGSGCYCGWEDKAGHGVCTGPKGGGYTGWS
HGFEVLGVYTPSGNTYQGTQAQKRGHIGLESKGWVYKGEWTHGFKRGYVRECA
NGAKYEGTWSNGLQDGYGTETSDG"

ORIGIN
Query Match 4.1%; Score 46.4; DB 9; Length 125020;
Best Local Similarity 10.8%; Pred. No. 0.61;
Matches 95; Conservative 397; Mismatches 378; Indels 6; Gaps 3;

Qy 108 AATCACCACTCAGAAACATTTAAGCGCTGTGCGCGCGCATTTCTGCGGTGAACGGTGT 167
Db 50582 RWTGKBGWSKSHSCHRRTWTDYRMKTYTCCCYTAIYMBBYHYCYDBCSGWSGWSGGW 50641

Qy 168 GACCATTTGTCAACGGTCCGAAAGGTTCTGGAATCGGAGCGCTTCGTTCTCTCGGAAAGGG 227
Db 50642 RGMKMBDWBMBBSMMGKSMGDTGDKDYWCYSSMSVDVVBWRWBDTSWGSBSHYMERGVH 50701

Qy 228 CTGGAATCTGCCCTCTGTCGCGGAAAGTGGTGCACATCGTGCAGACGAAATTTCTCA 287
Db 50702 TGYMRWVYGBCDTHVTYDGGHSGWGMKSCGYSDBGKCAKBSKMCCTGSMTKYTCYGY 50761

Qy 288 AGTTGTTGTCTATCACTCGGAGCTGGATCTGTTGGATGGTGCACATGACCATGCCCA 347
Db 50762 GKYWSKSAADBSYYSBSHYBWSYMCCKRBSTASSTDYRYRYKTYRVRSCSWYVK 50821

Qy 348 TACAGCTGTGTCAGCGCTCCA-CGACCTATGGAATGGAATCTTCGCGAGCTGGAAGAGCTG 406
Db 50822 YBSMSVSARKSGHKVRBBSMKCSRRKDMSTSSMWSRSMKDRKCSRCSHSHYSKMGKWS 50881

Qy 407 CGACCAACGGTCTCGTTCGCGTACGGAACGTAAGATGTACGGCTGCGACGATACGTGC 466
Db 50882 DVDRYACYSBMKYKHKYKDKDGBHSWGSKMKKGAWRVSRVSRMRCVSHDSMVD 50941

Qy 467 GCGCTGTGAAATCACTCACGGAAGAAACGGCTGGACGCTCCACGTTCCAGCCCTACTCA 526
Db 50942 KSRMRKGSAMRQMKCYCTYSMRSTKSRMSKMSWRWGSKC----YCYGYWCTKMRSM 50997

Qy 527 TGTTCAGTGTGATCGTGAAGAACATCTCGAATCTTCTCGGATCGGATCGATCTGCATC 586
Db 50998 GCSYSTGSSMKCMGWYCCWAGRSRYCKSGSRMSWAGSMASMKMTSGMSGRCC 51057
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 22:12:19 ; Search time 637 Seconds

(without alignments)
9394.568 Million cell updates/sec

Title: US-10-007-527A-1

Perfect score: 1140

Sequence: 1 atgaccagcgtaagtgtga.....taatatcggttcgaagtga 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1140	6	ABQ76122
C 2	1140	100.0	6334	6	ABQ76122 Rhodococc
C 3	1140	100.0	9652	6	ABQ76126 Plasmid p
C 4	1140	100.0	11241	6	ABQ76125 Plasmid p
C 5	1140	100.0	11241	12	ADH10182 E. coli-R
6	51.6	4.5	2297	2	AAV58945 B. breve
7	44.4	3.9	2000	8	ADA71938 Rice gene
8	40.6	3.6	1189	10	ACH00861 Murine PT
9	40.6	3.6	1244	10	ACH00860 Murine PT
10	40.6	3.6	1841	8	ACC00046 Mouse PTX
11	37.4	3.3	2708	8	ACC00048
12	37.4	3.3	4447	3	AAA90951
13	37.4	3.3	4447	3	AAA90934
14	37.4	3.3	4447	6	ABL49734
15	37.4	3.3	4447	6	ABL49733
16	37.4	3.3	4447	6	ADL22583
17	37.4	3.3	4447	6	ADL22582
18	37.4	3.3	4447	10	ADB66212
C 19	36.6	3.2	2000	8	ADA71779
C 20	36.4	3.2	2000	8	ADA71938
C 21	34.8	3.1	498	8	ACA18944

C 22	34.8	3.1	702	8	ACA14012
23	34.8	3.1	861	8	ACA53382
C 24	34.8	3.1	1631	3	AAA66016
C 25	34.8	3.1	4590	5	AHH24065
C 26	34.6	3.0	1505	2	AAQ55750
C 27	34.6	3.0	1954	6	ABT99802
C 28	34.4	3.0	1173	5	AAF26431
C 29	34.2	3.0	684	4	AAF26442
C 30	34.2	3.0	1314	6	ABQ90222
C 31	34.2	3.0	3399	3	ADC78561
C 32	34.2	3.0	3448	4	AAF26441
C 33	34.2	3.0	4938	4	AA559525
C 34	34.2	3.0	4938	8	ACF64454
C 35	34	3.0	534	6	ABQ41452
C 36	34	3.0	534	6	ABQ41452
C 37	34	3.0	1137	8	ACD05604
C 38	33.8	2.9	2019	12	ADQ22781
C 39	33.6	2.9	786	8	ACA00393
C 40	33.6	2.9	832	6	ABQ20710
C 41	33.6	2.9	832	6	ABQ20711
C 42	33.6	2.9	1134	3	AAA58791
C 43	33.6	2.9	1134	6	AAQ39776
C 44	33.6	2.9	1134	6	ABS53140
C 45	33.6	2.9	1281	8	ACA39926

ALIGNMENTS

RESULT 1

ABQ76122

ID ABQ76122 standard; DNA; 1140 BP.

XX

AC ABQ76122;

XX

DT 13-JAN-2003 (first entry)

XX

DE Rhodococcus AN12 replication protein Rep DNA.

XX

KW Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase; gene;
KW shuttle vector; Rep; db.

XX Rhodococcus erythropolis.

OS

XX WO200255709-A2.

XX

XX 18-JUL-2002.

PD

XX 12-DEC-2001; 2001WO-US047868.

PF

XX 12-DEC-2000; 2000US-0254868P.

PR

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA

XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

PI

XX WPI; 2002-557827/59.

XX

XX P-PSDB; ABB84278.

DR

XX New nucleic acid molecule encoding replication protein/plasmid stability

XX protein, useful in cloning and expression vectors, particularly shuttle

XX vectors for expression of heterologous genes in Rhodococcus species.

PS

XX Claim 2; Page 63-64; 96pp; English.

XX

CC This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile

CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence encodes the Rhodococcus AN12 Rep protein
CC described in the disclosure of the invention
XX
SQ Sequence 1140 BP; 241 A; 287 C; 347 G; 265 T; 0 U; 0 Other;
Query Match 100.0%; Score 1140; DB 6; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACACGCTGAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
DB 1 ATGACACGCTGAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
QY 61 TCCGATAGCGGGCATCCGGACGAACTGGACCAAACTTCAACAAATCACCACGTCA 120
DB 61 TCCGATAGCGGGCATCCGGACGAACTGGACCAAACTTCAACAAATCACCACGTCA 120
QY 121 GAAACATTAAACGCTGTGGCGCGGCGATTTCTGGCGTGAACGGTGTGACCAATTGTCAAC 180
DB 121 GAAACATTAAACGCTGTGGCGCGGCGATTTCTGGCGTGAACGGTGTGACCAATTGTCAAC 180
QY 181 GGTCCGAAAGTTCTGGATTCGGAGGCTTCTTCTTCTGGGAAAGGCTGGATTCGCCCC 240
DB 181 GGTCCGAAAGTTCTGGATTCGGAGGCTTCTTCTTCTGGGAAAGGCTGGATTCGCCCC 240
QY 241 TGCTGTGGGAAAGTGTGTCACATCGTCGACGAAATTTCTCAAGTTGTTCCTCAT 300
DB 241 TGCTGTGGGAAAGTGTGTCACATCGTCGACGAAATTTCTCAAGTTGTTCCTCAT 300
QY 301 CAACCTCGGACCTGGATCTGTTGCGATGTGTGACGATGACCATCGGCATACAGCTGGTCA 360
DB 301 CAACCTCGGACCTGGATCTGTTGCGATGTGTGACGATGACCATCGGCATACAGCTGGTCA 360
QY 361 CGGCTTCCACGACCTTAGACTGGAATTTCTGGCAGCTGTGAAAGCTGGACCAACGGTGT 420
DB 361 CGGCTTCCACGACCTTAGACTGGAATTTCTGGCAGCTGTGAAAGCTGGACCAACGGTGT 420
QY 421 CGTTGGCTGACGGAACGTAATGTACGCTCGGACGATAGTGGCGCTGTGAAATC 480
DB 421 CGTTGGCTGACGGAACGTAATGTACGCTCGGACGATAGTGGCGCTGTGAAATC 480
QY 481 ACTCACGGAACAAACGGCTGGCAGCTCCACGTTTCCAGCTTCTAGTGGTGCAT 540
DB 481 ACTCACGGAACAAACGGCTGGCAGCTCCACGTTTCCAGCTTCTAGTGGTGCAT 540
QY 541 GTGAGTGAACATCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGACTTCCAAA 600
DB 541 GTGAGTGAACATCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGACTTCCAAA 600
QY 601 CTCGTATCTCTGGATTTCTCGGCTCGGACCTACGTAATTCGGGTGCTCGATGACGAAAG 660
DB 601 CTCGTATCTCTGGATTTCTCGGCTCGGACCTACGTAATTCGGGTGCTCGATGACGAAAG 660
QY 661 ATCGGCGGTGAAGCTGATCAAGTTTCTCGGTGGTATCTGACGAAATTCGATCTGGCGTT 720
DB 661 ATCGGCGGTGAAGCTGATCAAGTTTCTCGGTGGTATCTGACGAAATTCGATCTGGCGTT 720
QY 721 GGTATGGAGGTTGGTATGGGACGGAAGAAAGTGTGACATGGGCAACCGTGCACCTGG 780
DB 721 GGTATGGAGGTTGGTATGGGACGGAAGAAAGTGTGACATGGGCAACCGTGCACCTGG 780
QY 781 GAAATCGCTGTGATGACGAGTGGGCGGAGATCCCAAGCGTTGGAATCTGGCGAGAAATTT 840
DB 781 GAAATCGCTGTGATGACGAGTGGGCGGAGATCCCAAGCGTTGGAATCTGGCGAGAAATTT 840
QY 841 GAGTTTGGTTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTGCGTCCGAGCT 900
DB 841 GAGTTTGGTTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTGCGTCCGAGCT 900

QY 901 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC 960
DB 901 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC 960
QY 961 ATGCTTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTGTGCCCTTACGTC 1020
DB 961 ATGCTTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTGTGCCCTTACGTC 1020
QY 1021 TTCGGCGAGATCTTCGACTCGTGAAGCTGGCGGAGCTTCGGAATAATCTTCGTGATCAC 1080
DB 1021 TTCGGCGAGATCTTCGACTCGTGAAGCTGGCGGAGCTTCGGAATAATCTTCGTGATCAC 1080
QY 1081 TTGCATTATCGATTTCCCGCGCGAGCGATGTGCGGCCCGGATTAATATCGTTTCGAAAGTGA 1140
DB 1081 TTGCATTATCGATTTCCCGCGCGAGCGATGTGCGGCCCGGATTAATATCGTTTCGAAAGTGA 1140

RESULT 2

ABQ76124/C
ID ABQ76124 standard; DNA; 6334 BP.

XX AC ABQ76124;
XX DT 13-JAN-2003 (first entry)
XX DE Rhodococcus AN12 derived plasmid pAN12 DNA.

XX Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector; circular; ds.

XX Rhodococcus erythropolis.

OS Synthetic.

XX WO200255709-A2.

XX 18-JUL-2002.

XX 12-DEC-2001; 2001WO-US047868.

XX 12-DEC-2000; 2000US-0254868P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

XX WPI; 2002-557827/59.

XX New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.

XX Claim 19; Page 68-71; 95pp; English.

XX This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence represents the Rhodococcus AN12 derived
CC plasmid pAN12 DNA described in the disclosure of the invention

XX SQ Sequence 6334 BP; 1386 A; 1807 C; 1681 G; 1460 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 6; Length 6334;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGACACGCTAAGTGTGAACACCTTTCCGCAAAAGACCGCCCTCCCGTCTCTCGTGTG	60
Db	3051	ATGACACGCTAAGTGTGAACACCTTTCCGCAAAAGACCGCCCTCCCGTCTCTCGTGTG	2992
QY	61	TCCGATAGCCGGCATCCGGCACGAACTCGGACCCAACTTCAACAATCACCACTCA	120
Db	2991	TCCGATAGCCGGCATCCGGCACGAACTCGGACCCAACTTCAACAATCACCACTCA	2932
QY	121	GAACATTTAACCCCTGTGGCGCGCGATTTCTGCGTGAACCGGTGTGACCAATGTCAAC	180
Db	2931	GAACATTTAACCCCTGTGGCGCGCGATTTCTGCGTGAACCGGTGTGACCAATGTCAAC	2872
QY	181	GGTCCGAAGGTTCTGGATTCCGAGCCCTTCGTTCTCGGAAAGGCTGGATCTGCCCC	240
Db	2871	GGTCCGAAGGTTCTGGATTCCGAGCCCTTCGTTCTCGGAAAGGCTGGATCTGCCCC	2812
QY	241	TGCTGTGCGGAAAGTCGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTCTCAT	300
Db	2811	TGCTGTGCGGAAAGTCGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTCTCAT	2752
QY	301	CAACTCGGGACTGGATCTGTTGCGATGTTGACATGCCGCATACAGCTGCTCAG	360
Db	2751	CAACTCGGGACTGGATCTGTTGCGATGTTGACATGCCGCATACAGCTGCTCAG	2692
QY	361	CGGCTCCACGACTTATGACCTGGACTTTTCGGCAGCTCGGAAAGCTGCCACCAACGGT	420
Db	2691	CGGCTCCACGACTTATGACCTGGACTTTTCGGCAGCTCGGAAAGCTGCCACCAACGGT	2632
QY	421	CGTTGGCGTACGGACGTTGAATGTACGGCTCGACCGGATACGTCGGCGCTGTTGAAATC	480
Db	2631	CGTTGGCGTACGGACGTTGAATGTACGGCTCGACCGGATACGTCGGCGCTGTTGAAATC	2572
QY	481	ACTCAGGAAAAACGGCTGGCAGCTCCACGTTCCACGGCTACTCATGTTCAAGTGTGAC	540
Db	2571	ACTCAGGAAAAACGGCTGGCAGCTCCACGTTCCACGGCTACTCATGTTCAAGTGTGAC	2512
QY	541	GTGAGTGAAACATCCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGATTTCCAAA	600
Db	2511	GTGAGTGAAACATCCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGATTTCCAAA	2452
QY	601	CTCGTATCTCGGATTTGCTCGGCACTACGTAATTCGGTGGTCTCGATGTACGAAG	660
Db	2451	CTCGTATCTCGGATTTGCTCGGCACTACGTAATTCGGTGGTCTCGATGTACGAAG	2392
QY	661	ATCGCGGTGAAGCTGATCAAGTTCTCGTGGCTATCTGACGAAATTCATCTGGCGTT	720
Db	2391	ATCGCGGTGAAGCTGATCAAGTTCTCGTGGCTATCTGACGAAATTCATCTGGCGTT	2332
QY	721	GGTATGGAGTTGTAGTGGCGACGGAAAAAGTGTGACATGCGCAACCGTGACCCCTGG	780
Db	2331	GGTATGGAGTTGTAGTGGCGACGGAAAAAGTGTGACATGCGCAACCGTGACCCCTGG	2272
QY	781	GAATCGTGTGTGATGCTAGTGGCGGGATCCACAGGTTGGAATGTGGGAGAAATTT	840
Db	2271	GAATCGTGTGTGATGCTAGTGGCGGGATCCACAGGTTGGAATGTGGGAGAAATTT	2212
QY	841	GAGTTTGGTTCGATGGGAGCTCGGCAATTCGGTGGTCCGTTGGATTCGTCGCCGAGCT	900
Db	2211	GAGTTTGGTTCGATGGGAGCTCGGCAATTCGGTGGTCCGTTGGATTCGTCGCCGAGCT	2152
QY	901	GGTCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGAAATTCGCCCGGTC	960
Db	2151	GGTCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGAAATTCGCCCGGTC	2092
QY	961	ATGTTTGGATCATTTCCGCGCGGATCGTGATGATTCGCACTTGTGGCGCTTACGTC	1020
Db	2091	ATGTTTGGATCATTTCCGCGCGGATCGTGATGATTCGCACTTGTGGCGCTTACGTC	2032
QY	1021	TTCCGGCGAGATCCTCGGACTCGTGAAGCTGGCGGACCTTGGGAAAAATCTTCGTGATCAC	1080

Db	2031	TTCCGGAGATCCTCGGACTCGTGAAAGCTGGCGGACTTGGAAAAATCTTCGTGATCAC	1972
QY	1081	TTGCATTAATTCGATTCGCCGACGCGATGTGGGCCCCGATTAATATCGGTTCCGAAGTCA	1140
Db	1971	TTGCATTAATTCGATTCGCCGACGCGATGTGGGCCCCGATTAATATCGGTTCCGAAGTCA	1912
RESULT 3			
ABQ76126/c			
ID	ABQ76126	standard; DNA; 9652 BP.	
XX	AC	ABQ76126;	
XX	AC	ABQ76126;	
DT	13-JAN-2003	(first entry)	
XX	DE	Plasmid pRHRI171 DNA.	
XX	DE	Plasmid pRHRI171 DNA.	
KW	KW	Plasmid stability protein; replication protein; ethylene forming enzyme;	
KW	KW	carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;	
KW	KW	polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;	
KW	KW	alcohol dehydrogenase; terpene synthase; cholesterol oxidase;	
KW	KW	shuttle vector; circular; ds.	
XX	OS	Synthetic.	
XX	XX		
PN	WO200255709-A2.		
XX	XX		
PD	18-JUL-2002.		
XX	XX		
PF	12-DEC-2001; 2001WO-US047868.		
XX	XX		
PR	12-DEC-2000; 2000US-0254868P.		
XX	XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E I.		
XX	XX		
PI	Bramucci MG, Cheng Q, Kostichka KN, Tomb J;		
XX	XX		
DR	WPI; 2002-557827/59.		
XX	XX		
PT	New nucleic acid molecule encoding replication protein/plasmid stability		
PT	protein, useful in cloning and expression vectors, particularly shuttle		
PT	vectors for expression of heterologous genes in Rhodococcus species.		
XX	XX		
PS	Claim 27; Page 72; 96pp; English.		
XX	XX		
CC	This invention describes a novel nucleic acid encoding a replication		
CC	protein or a plasmid stability protein. The product of the invention is		
CC	useful for expression of nucleic acid such as genes encoding enzymes		
CC	involved in the production of isoprenoid molecules, polyhydroxyalkanoic		
CC	acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile		
CC	hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol		
CC	dehydrogenase, terpene synthases, and cholesterol oxidase in an		
CC	Actinomycetales bacteria. The replication protein or plasmid stability		
CC	protein are useful in cloning and expression vectors and particularly in		
CC	shuttle vectors for the expression of homologous and heterologous genes		
CC	in Rhodococcus sp. This sequence represents the Plasmid pRHRI171 DNA		
CC	described in the disclosure of the invention		
XX	XX		
SQ	Sequence 9652 BP; 2097 A; 2729 C; 2559 G; 2267 T; 0 U; 0 Other;		
Query Match 100.0%; Score 1140; DB 6; Length 9652;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGACACGCTAAGTGTGAACACCTTTCCGCAAAAGACCGCCCTCCCGTCTCTCGTGTG	60
Db	5052	ATGACACGCTAAGTGTGAACACCTTTCCGCAAAAGACCGCCCTCCCGTCTCTCGTGTG	4993
QY	61	TCCGATAGCCGGCATCCGGCACGAACTCGGACCCAACTTCAACAATCACCACTCA	120
Db	4992	TCCGATAGCCGGCATCCGGCACGAACTCGGACCCAACTTCAACAATCACCACTCA	4933
QY	121	GAACATTTAAGCCTGTGGCGCGGCGGATTTCTGGCGTGAACGGTGTGACCATGTGCAAC	180

```
Db 4932 GAAACATTTAAAGCCTTGCGCGCGGATTTCTGGGTGAACGGTGTGACCAATGTCAAC 4873
QY 181 GGTCCGAAAGTTCTGGATTCGAGAGCCTTCTGTTCTCGGAAAGGGCTGATCTGCCCC 240
Db 4872 GGTCCGAAAGTTCTGGATTCGAGAGCCTTCTGTTCTCGGAAAGGGCTGATCTGCCCC 4813
QY 241 TCGTGTGGGGAAGTGGTGCACATCTGCAGAGAAATTTCTCAAGTTGTGCTCAT 300
Db 4812 TCGTGTGGGGAAGTGGTGCACATCTGCAGAGAAATTTCTCAAGTTGTGCTCAT 4753
QY 301 CAACTCGGAGTGGATCTGTTGGATGGTGAACATGACCATGCGCATACAGTGTGTAG 360
Db 4752 CAACTCGGAGTGGATCTGTTGGATGGTGAACATGACCATGCGCATACAGTGTGTAG 4693
QY 361 CGGCTCCAGCACTATGGAAGTGGATCTGCGAGAGCTGGAAGCTGCGACCAACAGTGT 420
Db 4692 CGGCTCCAGCACTATGGAAGTGGATCTGCGAGAGCTGGAAGCTGCGACCAACAGTGT 4633
QY 421 CGTTCGGCTACGGAAGTGAATGTACGGCTGCGAGGATAGTGGCGGCTGTTGAAATC 480
Db 4632 CGTTCGGCTACGGAAGTGAATGTACGGCTGCGAGGATAGTGGCGGCTGTTGAAATC 4573
QY 481 ACTCAGGAAAAAAGCGCTGGCACTGCTCCAGCTTTCAGCGCTACTCATGTTCAAGTGTGAC 540
Db 4572 ACTCAGGAAAAAAGCGCTGGCACTGCTCCAGCTTTCAGCGCTACTCATGTTCAAGTGTGAC 4513
QY 541 GTGAGTGAAGCAATCTCGAATCTTCTCGAATGCGAATGTCGATGCGTGGACTTCCAAA 600
Db 4512 GTGAGTGAAGCAATCTCGAATCTTCTCGAATGCGAATGTCGATGCGTGGACTTCCAAA 4453
QY 601 CTCGTATCTCTGGATTTCTCGCGCACTACGTAATTCGGTGGTCTCGATCTAGAGAG 660
Db 4452 CTCGTATCTCTGGATTTCTCGCGCACTACGTAATTCGGTGGTCTCGATCTAGAGAG 4393
QY 661 ATCGCGGGTGAAGCTGATCAAGTTCTCGCTGGTATCTCAGCAAAATTCATCTGCGGTT 720
Db 4392 ATCGCGGGTGAAGCTGATCAAGTTCTCGCTGGTATCTCAGCAAAATTCATCTGCGGTT 4333
QY 721 GGTATGAGAGTTGGTAGTGGAGCGAAGAAAGTGGTGCAGTGGCAACCGTGACCCCTGG 780
Db 4332 GGTATGAGAGTTGGTAGTGGAGCGAAGAAAGTGGTGCAGTGGCAACCGTGACCCCTGG 4273
QY 781 GAAATCGCTGTTGATGACGTGGCGGGGATCCACAGCGTGGAGCTGGCGAGATTT 840
Db 4272 GAAATCGCTGTTGATGACGTGGCGGGGATCCACAGCGTGGAGCTGGCGAGATTT 4213
QY 841 GAGTTTGGTTTCGATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTGCGTCCCGAGCT 900
Db 4212 GAGTTTGGTTTCGATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTGCGTCCCGAGCT 4153
QY 901 GGTCTTGGGGAGAACTAACAGATGCTCAGATGCTGTTGAGCAGAGAAATCTGCCCGCGTC 960
Db 4152 GGTCTTGGGGAGAACTAACAGATGCTCAGATGCTGTTGAGCAGAGAAATCTGCCCGCGTC 4093
QY 961 ATGTTTGGCATCATTCGCGCGGATCGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Db 4092 ATGTTTGGCATCATTCGCGCGGATCGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 4033
QY 1021 TTCGCGGAGATCTCGGACTCGTTCGAAAGTGGCGGACTTGGGAAATCTTCGTGATCAC 1080
Db 4032 TTCGCGGAGATCTCGGACTCGTTCGAAAGTGGCGGACTTGGGAAATCTTCGTGATCAC 3973
QY 1081 TTGCATTATCGATTGCCCGAGGGATGTGCGGCCCCCGCATATATCGGTTGCGAAGTGA 1140
Db 3972 TTGCATTATCGATTGCCCGAGCGGATGTGCGGCCCCCGCATATATCGGTTGCGAAGTGA 3913
```

RESULT 4

AB076125/c

ID AB076125 standard; DNA; 11241 BP.

XX

AC AB076125;

```
XX 13-JAN-2003 (first entry)
DT Plasmid pRHBR17 DNA.
XX
XX Plasmid stability protein; replication protein; ethylene forming enzyme;
XX carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
XX polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
XX alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
XX shuttle vector; circular; ds.
XX
OS Synthetic.
XX
XX WO200255709-A2.
XX
XX 18-JUL-2002.
XX
XX 12-DEC-2001; 2001WO-US047868.
XX
XX 12-DEC-2000; 2000US-0254868P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX WPI; 2002-557827/59.
XX
XX New nucleic acid molecule encoding replication protein/plasmid stability
XX protein, useful in cloning and expression vectors, particularly shuttle
XX vectors for expression of heterologous genes in Rhodococcus species.
XX
XX Claim 26; Page 71-72; 96pp; English.
XX
XX This invention describes a novel nucleic acid encoding a replication
XX protein or a plasmid stability protein. The product of the invention is
XX useful for expression of nucleic acid such as genes encoding enzymes
XX involved in the production of isoprenoid molecules, polyhydroxyalkanoic
XX acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
XX hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
XX dehydrogenase, terpene synthases, and cholesterol oxidase in an
XX Actinomycetales bacteria. The replication protein or plasmid stability
XX protein are useful in cloning and expression vectors and particularly in
XX shuttle vectors for the expression of homologous and heterologous genes
XX in Rhodococcus sp. This sequence represents the Plasmid pRHBR17 DNA
XX described in the disclosure of the invention
XX
SQ Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;
Query Match 100.0%; Score 1140; DB 6; Length 11241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACACAGCTAAGTGTGAACACCTTTCCGCAAGACCGGCTCCCTCGTGTGCG 60
Db 6641 ATGACACAGCTAAGTGTGAACACCTTTCCGCAAGACCGGCTCCCTCGTGTGCG 6582
QY 61 TCCGATAGCGCGGATCGGACCACTGCGACCAACTTCACAAATCACCAGTCA 120
Db 6581 TCCGATAGCGCGGATCGGACCACTGCGACCAACTTCACAAATCACCAGTCA 6522
QY 121 GAAACATTTAAAGCCTTGCGCGGCGGATTTCTGGCGTGAACGGTGTGACCAATGTCAAC 180
Db 6521 GAAACATTTAAAGCCTTGCGCGGCGGATTTCTGGCGTGAACGGTGTGACCAATGTCAAC 6462
QY 181 GGTCCGAAAGTTCTGGATTCGAGAGCCTTCTGTTCTCGGAAAGGGCTGATCTGCCCC 240
Db 6461 GGTCCGAAAGTTCTGGATTCGAGAGCCTTCTGTTCTCGGAAAGGGCTGATCTGCCCC 6402
QY 241 TCGTGTGGGGAAGTGGTGCACATCTGCAGAGAAATTTCTCAAGTTGTGCTCAT 300
Db 6401 TCGTGTGGGGAAGTGGTGCACATCTGCAGAGAAATTTCTCAAGTTGTGCTCAT 6342
QY 301 CAACTCGGAGTGGATCTGTTGGATGGTGAACATGACCATGCGCATACAGTGTGTAG 360
```



```
Db 6341 CAATCGGACTGGATCTGTTGGATGTTGACGATGACATCGGCATACAGCTGGTCAG 6282
Qy 361 CGGCTCCACGACCTATGAGCTGGAATTTCCGAGCCTGGAAGCTGCGACCAACGGTCGT 420
Db 6281 CGGCTCCACGACCTATGAGCTGGAATTTCCGAGCCTGGAAGCTGCGACCAACGGTCGT 6222
Qy 421 CGTTGGCGTACGGACGTTGAATGTACGGCTGCGAGCGATACGTTGGCGCTGTTGAATC 480
Db 6221 CGTTGGCGTACGGACGTTGAATGTACGGCTGCGAGCGATACGTTGGCGCTGTTGAATC 6162
Qy 481 ACTCAGGAAAAAAGCGCTGGCAGCTCCACGCTTACGCGCTACTCATGTTCAAGTGGTGCAC 540
Db 6161 ACTCAGGAAAAAAGCGCTGGCAGCTCCACGCTTACGCGCTACTCATGTTCAAGTGGTGCAC 6102
Qy 541 GTGAGTGAGAACATCTCGAATCTTTCGAGATCGCATGTTTCGATCGGTGGACTTCCAAA 600
Db 6101 GTGAGTGAGAACATCTCGAATCTTTCGAGATCGCATGTTTCGATCGGTGGACTTCCAAA 6042
Qy 601 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTTGGGTGGTCTCGATGTAGGAAAG 660
Db 6041 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTTGGGTGGTCTCGATGTAGGAAAG 5982
Qy 661 ATCGCGGTGAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAAAATTCGATCTGGCGTT 720
Db 5981 ATCGCGGTGAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAAAATTCGATCTGGCGTT 5922
Qy 721 GGTATGAGGTTGTGATGGGACGGAAGAAAGTGTGACATGCGCAACCGTGCACCTGG 780
Db 5921 GGTATGAGGTTGTGATGGGACGGAAGAAAGTGTGACATGCGCAACCGTGCACCTGG 5862
Qy 781 GAAATCGCTGTGATGAGTGGCGGGATCCACAGGTTTGGAACTGTGCGGAGATTT 840
Db 5861 GAAATCGCTGTGATGAGTGGCGGGATCCACAGGTTTGGAACTGTGCGGAGATTT 5802
Qy 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGCTCCGTTGAGTTCGCTGCCGAGCT 900
Db 5801 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGCTCCGTTGAGTTCGCTGCCGAGCT 5742
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 5682
Qy 961 ATGTTGCGATCANTCCGGCGGATCGTGGATGATGATTCGGAATTTGCGGCTTACGTC 1020
Db 5681 ATGTTGCGATCANTCCGGCGGATCGTGGATGATGATTCGGAATTTGCGGCTTACGTC 5622
Qy 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC 1080
Db 5621 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC 5562
Qy 1081 TTGCATATCGATTGCCGCGAGGATGTGGGCCCCCGATTAATATCGGTTTCGCAAGTGA 1140
Db 5561 TTGCATATCGATTGCCGCGAGGATGTGGGCCCCCGATTAATATCGGTTTCGCAAGTGA 5502
```

RESULT 5

ADH10182/c
ID ADH10182 standard; DNA; 11241 BP.

XX AC ADH10182;

XX AC ADH10182;

DT 11-MAR-2004 (first entry)

XX DE E. coli-Rhodococcus shuttle plasmid pRHB17.

XX DE Aryl carotenoid; beta-ionone; carotene desaturase; CrtU; isorenieratene;
KW chlorobactene; pharmaceutical; food supplement; animal feed additive;
KW food colourant; cosmetic; crt0; ds.

XX OS Synthetic.

XX PN WO2003093200-A2.

XX

PD 13-NOV-2003.

XX PF 06-MAY-2003; 2003WO-US014111.

XX PR 06-MAY-2002; 2002US-0378312P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Cheng Q, Tao L, Rouviere PE;

XX PS WPI; 2004-081862/08.

XX PT Production of aryl carotenoids useful in e.g. pharmaceuticals and food

XX PT colorants involves bioconversion of cyclic carotenoids comprising beta-

XX PS ionone ring in the presence of carotene desaturase.

XX PS Example 3; SEQ ID NO 13; 66pp; English.

XX CC The invention relates to production of aryl carotenoid compounds. The

XX CC method involves transforming a host cell comprising a cyclic carotenoid

XX CC having at least one beta-ionone ring with a gene encoding a carotene

XX CC desaturase (crtU) followed by growing the transformed host cell. Also

XX CC provided is a method for regulating aryl carotenoid biosynthesis in a

XX CC host cell. The method is useful in the production of aryl carotenoids

XX CC such as isorenieratene and chlorobactene, which are useful in

XX CC pharmaceuticals, food supplements, animal feed additives, food colorants

XX CC and cosmetics. The present sequence represents a E. coli-Rhodococcus

XX CC shuttle plasmid pRHB17 nucleotide sequence

XX SQ Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 12; Length 11241;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACCAGCTAAGTGTGAACACCTTTCGGGAAAAGACCGGCTCCCGTCTCTCGTGTG 60

Db 6641 ATGACCAGCTAAGTGTGAACACCTTTCGGGAAAAGACCGGCTCCCGTCTCTCGTGTG 6582

Qy 61 TCCGATAGCGCGCATCCGCGACGAACTCGCGACCCAACTTCAACAAATCACCACGTC 120

Db 6581 TCCGATAGCGCGCATCCGCGACGAACTCGCGACCCAACTTCAACAAATCACCACGTC 6522

Qy 121 GAAACATTTAACGCTGTGGCCGCGATTTCTGGCGTGAAACGGTGTGACCAATTTGTCAAC 180

Db 6521 GAAACATTTAACGCTGTGGCCGCGATTTCTGGCGTGAAACGGTGTGACCAATTTGTCAAC 6462

Qy 181 GGTCCGAAAGGTTCTGGATTCGGAGGCTTCTGTTCTCGGAAAAGGGCTGGATCTGCCCC 240

Db 6461 GGTCCGAAAGGTTCTGGATTCGGAGGCTTCTGTTCTCGGAAAAGGGCTGGATCTGCCCC 6402

Qy 241 TGCTGTGGGAAAAGTTCGGTGACATCTGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 300

Db 6401 TGCTGTGGGAAAAGTTCGGTGACATCTGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 301 CAATCGGACTGGATCTGTTGCGATGTTGACGATGACCAATCGGCCATACAGCTGGTCA 360

Db 6341 CAATCGGACTGGATCTGTTGCGATGTTGACGATGACCAATCGGCCATACAGCTGGTCA 6282

Qy 361 CGGCTCCACGACCTATGAGCTGGAATTTCCGAGCCTTGGAAAGCTGGACCAACGGTCGT 420

Db 6281 CGGCTCCACGACCTATGAGCTGGAATTTCCGAGCCTTGGAAAGCTGGACCAACGGTCGT 6222

Qy 421 CGTTGGCGTACGGAAACGTTGAATGTACGGCTGCGACGATACGTCGCGCTGTTGAAATC 480

Db 6221 CGTTGGCGTACGGAAACGTTGAATGTACGGCTGCGACGATACGTCGCGCTGTTGAAATC 6162

Qy 481 ACTCAGGAAAAAAGCGCTGGCAGCTCCACGTTTCAGCGCTACTCATGTTTCAGTGGTGC 540

Db 6161 ACTCAGGAAAAAAGCGCTGGCAGCTCCACGTTTCAGCGCTACTCATGTTTCAGTGGTGC 6102

Qy 541 GTGAGTGAGAACATCTCGAATCTTTCGAGTTCGATGTTTCGATCGGTGAGCAATTC 600


```

Db 6101 GTGAGTGAGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGGACTTCCAAA 6042
Qy 601 CTCGTATCTCTGGGATTTGCTGCGCCACTAGCTAATTCGGGTGGTCTCGATGTACGAAAG 660
Db 6041 CTCGTATCTCTGGGATTTGCTGCGCCACTAGCTAATTCGGGTGGTCTCGATGTACGAAAG 5982
Qy 661 ATCGGCGGTGAAGCTGATCAAGTTCTCGTGGGTATCTGACGAAAATTTGCATCTGGCGTT 720
Db 5981 ATCGGCGGTGAAGCTGATCAAGTTCTCGTGGGTATCTGACGAAAATTTGCATCTGGCGTT 5922
Qy 721 GGTATGAGGTTGGTGGGACGAAAGAGTGTGCACATGGCAACCGTGCACCTGG 780
Db 5921 GGTATGAGGTTGGTGGGACGAAAGAGTGTGCACATGGCAACCGTGCACCTGG 5862
Qy 781 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAAGCGTTGGAACTGTGGCGAATTT 840
Db 5861 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAAGCGTTGGAACTGTGGCGAATTT 5802
Qy 841 GAGTTGTTGATCGGAGCTCGGCAATCGCGTGGTCCGTTGGATTTGGTCCCGAGCT 900
Db 5801 GAGTTGTTGATCGGAGCTCGGCAATCGCGTGGTCCGTTGGATTTGGTCCCGAGCT 5742
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGAACTCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGAACTCTGCCCGGTC 5682
Qy 961 ATGTTTGGCATCATTCGCGCGCATCGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Db 5681 ATGTTTGGCATCATTCGCGCGCATCGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 5622
Qy 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGCATTCGGGAAATCTTCGTGATCAC 1080
Db 5621 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGCATTCGGGAAATCTTCGTGATCAC 5562
Qy 1081 TTGCATTATCGATTGCCGCGAGGATGTGCGGCCCCGATAATATCGGTTTCGCAAGTGA 1140
Db 5561 TTGCATTATCGATTGCCGCGAGGATGTGCGGCCCCGATAATATCGGTTTCGCAAGTGA 5502

```

RESULT 6

AAV58945
ID AAV58945 standard; DNA; 2297 BP.

XX AAV58945;

AC AAV58945;

XX 05-JAN-1999 (first entry)

XX B. breve essential region gene.

XX Essential region gene; shuttle vector; ds.

XX Bifidobacterium breve.

Key	Location/Qualifiers
PH -35_signal	387..392
FT	/tag= a
FT -10_signal	410..415
FT	/tag= b
FT RBS	525..530
FT	/tag= c
FT CDS	539..2023
FT	/tag= d

XX JP10262670-A.

XX 06-OCT-1998.

XX 27-MAR-1997; 97JP-00091387.

XX 27-MAR-1997; 97JP-00091387.

XX (HONS) YAKULT HONSHA KK.

XX WPI; 1998-587288/50.
DR P-PSDB; AAW73071.

XX Shuttle vector for a Bifidobacterium species - contains specific elements from existing plasmids e.g. pNBb1 of ATCC 15698.

XX Disclosure; Page 14-15; 17pp; Japanese.

XX This sequence represents the essential region gene from plasmid pNBb1 used in the vector of the invention. The vector is a shuttle vector for a Bifidobacterium, and contains: (a) a replicated essential region originated from a plasmid pNBb1 of Bifidobacterium breve ATCC 15698; (b) a replicated essential region originated from a plasmid of E. coli; and (c) an antibiotic-resistant gene originated from a plasmid of E. coli and an antibiotic-resistant gene functioning by a Bifidobacterium. The shuttle vector can be used to transform various Bifidobacteria species

Qy Sequence 2297 BP; 524 A; 659 C; 689 G; 425 T; 0 U; 0 Other;
Query Match 4.5%; Score 51.6; DB 2; Length 2297;
Best Local Similarity 52.9%; Fred. No. 3.5e-05;
Matches 111; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy	394	GCCTGGAAGCTGCGACCAACGCTGCTGTTGGGCTACGGAACGTTGAAATGTACGGCTGC	453
Db	1184	GGCTGGACAGATGATTAAAGGAGCCCTTGGCAACGGGCTCGAAGCTTGGAAATC	1243
Qy	454	GACGATACGTCGCGCTGTTGAAATCACTCAAGAAAAACGGCTGCACGTTCCACGTT	513
Db	1244	AGGGGTTTCTGTCGCGGCTGTTGAAATCACTACGTTGAAACGGCTGCACGTTCCACATT	1303
Qy	514	CACGGCTACTCATGTTTCAGTGTGACGTGAGTGAACATCTCTGAAATCTCTTCGGAT	573
Db	1304	CATTTGTCATGTTTCTCGATGGGCTGTCGACGATGCGGCTGAGCAATGCGCAA	1363
Qy	574	GCGATGTTGATCGGTGCGACTTCCAAATC	603
Db	1364	TGGCTGCTCGATCGCTGGAACCATGTC	1393

RESULT 7

ADA71938

ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to

KW vaccine; cancer; cytostatic; gene; ss.
 XX
 OS Mus sp.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 72..1235
 FT /*tag= a
 FT /product= "modified PTX3"
 XX
 PN WO2003072603-A2.
 XX
 PD 04-SEP-2003.
 XX
 PF 25-FEB-2003; 2003WO-IT000104.
 XX
 PR 28-FEB-2002; 2002IT-RM000109.
 XX
 PA (SIGT) SIGMA-TAU IND FARM RIUNITE SPA.
 XX
 PI De Santis R, Salvatori G;
 XX
 DR WPI; 2003-731602/69.
 DR P-PSDB; ABG75068.
 XX
 PT New derivatives of human or murine pentraxin-3, useful for preparing
 PT autologous vaccines for treating tumors, also new nucleic acid encoding
 PT them.
 XX
 PS Claim 7; Page 31; 32pp; English.
 XX
 CC The present invention relates to derivatives of murine and human
 CC pentraxin-3 (PTX3) (shown in ABG75068-ABG75073). These sequences, and
 CC biotinylated derivatives of PTX3, are used to make autologous vaccines,
 CC based on inactivated tumour cells, to treat solid or haematological
 CC tumours. The present sequence is a derivative of the murine PTX3 cDNA
 CC shown in the exemplification of the invention
 XX
 SQ Sequence 1244 BP; 261 A; 320 C; 405 G; 258 T; 0 U; 0 Other;
 Query Match 3.6%; Score 40.6; DB 10; Length 1244;
 Best Local Similarity 48.5%; Pred. No. 0.083;
 Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
 QY 696 TCTGACGAAATTGTCATCTGGCGTGGTATGAGAGTTGGTAGTCGCGACGCGGAAAGTGG 755
 Db 331 TCCGTGGAGAGCTGCAGCGGCTGCGGCGCAGAGCTGGGCGGCGCATGGCGA 390
 QY 756 TCGACATGGCAACCTGCACCTGGGAAATCGCTGTGATGAGTGGGCGGGATCCACA 815
 Db 391 GGCCGTGGCAGCCGCTGCGCCGCGAGACCGCAGAGCTGGTGGGGCGCTGGAGCCGCTGC 450
 QY 816 AGCGTTGGAACCTGTGGCGAGAAATTGAGTTTGGTTTCGATGGGACGTCGCGGCAATCGCGTG 875
 Db 451 TGCAGGAGAGCGCTGACGCGAGCCTCAGGCTGGCGCGCTGAGGACGCGGAGGCGCGC 510
 QY 876 GTCCCGTGGATTGCGTCCCGAGCTGGTTTGGGCGCAGACTAACAGATGC 926
 Db 511 GACCCGAGCGCAGTGGCTGGCCCTAGGCGCTGTGCTGAGGAACTGCGGC 561
 RESULT 10
 ACC00046
 ID ACC00046 standard; cDNA; 1841 BP.
 XX
 AC ACC00046;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Mouse PTX3 nucleotide sequence.
 XX
 KW Mouse pentraxin 3; PTX3; reproductive ability; antiinfertility;
 KW infertility; ss.

XX Mus sp.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 144..1286
 FT /*tag= a
 FT /product= "PTX3 mouse"
 XX
 PN WO2003011326-A1.
 XX
 PD 13-FEB-2003.
 XX
 PF 18-JUL-2002; 2002WO-IT000473.
 XX
 PR 03-AUG-2001; 2001US-0309472P.
 XX
 PA (SIGT) SIGMA-TAU IND FARM RIUNITE SPA.
 XX
 PI Mantovani A;
 XX
 DR WPI; 2003-239472/23.
 DR P-PSDB; ABR58222.
 XX
 PT Use of recombinant human pentraxin 3 gene and protein for preparing a
 PT medicament for increasing the reproductive ability in a female subject,
 PT or for diagnosing or treating female infertility.
 XX
 PS Disclosure; Page 54; 63pp; English.
 XX
 CC The present invention relates to the use of recombinant human pentraxin 3
 CC (PTX3) for preparing a medicament for increasing the reproductive ability
 CC in a female subject, used as an antiinfertility drug. The virals or
 CC plasmid vectors containing the human PTX3 cDNA are useful for treating
 CC female subjects in need of increasing the reproductive ability. The PTX3
 CC protein is useful as a diagnostic marker of the reproductive ability in
 CC human female, and PTX3 is useful as a target protein for the screening of
 CC pharmaceutical compounds to assess their capability to affect the
 CC reproductive ability in a female subject. The present sequence represents
 CC mouse PTX3 nucleotide sequence
 XX
 SQ Sequence 1841 BP; 444 A; 405 C; 527 G; 465 T; 0 U; 0 Other;
 Query Match 3.6%; Score 40.6; DB 8; Length 1841;
 Best Local Similarity 48.5%; Pred. No. 0.1;
 Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
 QY 696 TCTGACGAAATTGTCATCTGGCGTGGTATGAGAGTTGGTAGTCGCGACGCGGAAAGTGG 755
 Db 341 TCCGTGGAGAGCTGCAGCGGCTGCGGCGCAGAGCTGGGCGGCGCATGGCGA 400
 QY 756 TCGACATGGCAACCTGCACCTGGGAAATCGCTGTGATGAGTGGGCGGGATCCACA 815
 Db 401 GGCCGTGGCAGCCGCTGCGCCCGCAGACCGCAGAGCTGGTGGGGCGCTGGAGCCGCTGC 460
 QY 816 AGCGTTGGAACCTGTGGCGAGAAATTGAGTTTGGTTTCGATGGGACGTCGCGGCAATCGCGTG 875
 Db 461 TGCAGGAGAGCGCTGACGCGAGCCTCAGGCTGGCGCGCTGAGGACGCGGAGGCGCGC 520
 QY 876 GTCCCGTGGATTGCGTCCCGAGCTGGTTTGGGCGCAGACTAACAGATGC 926
 Db 521 GACCCGAGCGCAGTGGCTGGCCCTAGGCGCTGTGCTGAGGAACTGCGGC 571
 RESULT 11
 ACC00048
 ID ACC00048 standard; cDNA; 2708 BP.
 XX
 AC ACC00048;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Mouse upstream regulatory region.
 XX

KW Mouse pentraxin 3; PTX3; reproductive ability; antiinfertility;
 XX infertility; ss.
 OS Mus sp.

XX WO2003011326-A1.
 XX 13-FEB-2003.

XX 18-JUL-2002; 2002WO-IT000473.
 XX 03-AUG-2001; 2001US-0309472P.

XX (SIGMA) SIGMA-TAU IND FARM RIUNITE SPA.
 XX Mantovani A;

XX WPI; 2003-239472/23.
 XX Use of recombinant human pentraxin 3 gene and protein for preparing a
 PT medicament for increasing the reproductive ability in a female subject,
 PT or for diagnosing or treating female infertility.

XX Disclosure; Page 57-58; 63pp; English.
 XX The present invention relates to the use of recombinant human pentraxin 3
 CC (PTX3) for preparing a medicament for increasing the reproductive ability
 CC in a female subject, used as an antiinfertility drug. The virals or
 CC plasmid vectors containing the human PTX3 cDNA are useful for treating
 CC female subjects in need of increasing the reproductive ability. The PTX3
 CC protein is useful as a diagnostic marker of the reproductive ability in
 CC human female, and PTX3 is useful as a target protein for the screening of
 CC pharmaceutical compounds to assess their capability to affect the
 CC reproductive ability in a female subject. The present sequence represents
 CC mouse PTX3 upstream regulatory region nucleotide sequence

XX SQ Sequence 2708 BP; 667 A; 704 C; 640 G; 697 T; 0 U; 0 Other;
 Query Match 3.3%; Score 37.4; DB 8; Length 2708;
 Best Local Similarity 47.6%; Pred. No. 1.3;
 Matches 110; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
 QY 696 TCTGACGAAATTCATCTGGCGTTGGTATGGAGGTTGGTAGTGGCGGCGGCGGCGCATGGCGA 755
 DB 2198 TCCGTGGAGAGCTGACAGCGGTCGGGTGAGAGTGGCGGCGGCGGCGGCGCATGGCGA 2257
 QY 756 TCGACATGGCAACCGTGCACCTCGGAAATTCGTTGATGAGTGGCGGCGGATCCACA 815
 DB 2258 GGCGGTGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2317
 QY 816 AGCGTTGGAAGTGGCGGAGAGTTGAGTTGGTTCGATGAGGAGCGTGGGCAATCGCGTG 875
 DB 2318 TGCAGGAGCGGTGACCGGAGCTTCAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2377
 QY 876 GTCCCGTGGATGGCTGGCGGAGCTGCTTGGGCGGAGAACTAACAGATGC 926
 DB 2378 GACCCGAGGCGACAGTGCCTGGCGCTAGGCGCTGTGCTGGAGGAACTGGCGG 2428

RESULT 12
 AAA90951
 ID AAA90951 standard; DNA; 4447 BP.

XX AAA90951;
 XX 15-SEP-2003 (revised)
 DT 15-JAN-2001 (first entry)

XX B. lactofermentum pSPK6 coding sequence.

XX Temperature sensitive plasmid; TSPCR; protein production;
 KW temperature sensitive replication control region; pSPK6; ds.
 XX

OS Corynebacterium glutamicum.
 XX Key Location/Qualifiers
 FH 1318..2601
 FT /*tag= a
 XX

XX EF1038966-A1.

XX 27-SEP-2000.

XX 16-MAR-2000; 2000EP-00105326.

XX 16-MAR-1999; 99JP-00069896.

XX (AJIN) AJINOMOTO CO INC.

XX Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;

XX WPI; 2000-573832/54.

XX P-PSDB; AAY97537.

XX Plasmids containing a temperature sensitive replication control regions
 PT useful for breeding microorganisms for the production of amino acids by
 PT fermentation.

XX Disclosure; Page 19-22; 29pp; English.

XX This sequence encodes the Brevibacterium lactofermentum pSPK6 protein.
 CC The invention relates to a plasmid containing a temperature sensitive
 CC replication control region (TSPCR) and a marker gene (MG). The TSPCR is
 CC derived from plasmid pM330 harboured by Brevibacterium lactofermentum
 CC (ATCC 13869) and allows the plasmid to replicate autonomously at a low
 CC temperature but does not allow the plasmid to replicate autonomously at
 CC an elevated temperature in coryneform bacteria within a temperature range
 CC in which the bacteria can grow. The plasmid can be used for modifying a
 CC chromosomal gene in a coryneform bacterium, which may be used for the
 CC production of useful substances, such as amino acids, by fermentation, to
 CC change their genetic traits. Therefore, the plasmid can be used for
 CC breeding microorganisms for the production of amino acids by
 CC fermentation. The plasmid comprises a TSPCR that allows the plasmid to
 CC replicate autonomously at an elevated temperature in coryneform bacteria
 CC within a temperature range in which the bacteria can grow. (Updated on 15
 CC -SEP-2003 to standardise OS field)

XX SQ Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;

Query Match 3.3%; Score 37.4; DB 3; Length 4447;
 Best Local Similarity 53.0%; Pred. No. 1.7;
 Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 493 AACGGCTGGCAGTCCACGTTACGCGCTACTCATGTTTCAGTGGTGGTGGAGTGAAC 552
 DB 1795 AACGGTGGCAGTTCGACCGCAACATGCTGTTGTTCTTGGATCGTCCACGTCTGACGAT 1854

QY 553 ATCTCCGAATCTTCTCGGATCGGATGCGATGTTGATCGGTGGAGTCCAACTGATCTCTG 612
 DB 1855 GAACCTAAGGCATTTGAGGATTCATGTTTCCCGCTGGTCTGCTGGTGGTTAGGCC 1914

QY 613 GGATTTCGCGCCACTACGTAATTCGGGTG 643
 DB 1915 GGATGACCGCGCATCTGCGTGAGCAGCGGG 1945

RESULT 13

AAA90934
 ID AAA90934 standard; DNA; 4447 BP.

XX AAA90934;

XX 15-SEP-2003 (revised)

DT 15-JAN-2001 (first entry)

XX B. lactofermentum p48K coding sequence.

XX	Temperature sensitive plasmid; TSRCR; protein production; temperature sensitive replication control region; p48K; ds.	XX	29-AUG-2003 (revised)
XX		XX	29-MAY-2002 (first entry)
KW		DE	Brevibacterium lactofermentum p48K nucleotide sequence SEQ ID NO:7.
XX		XX	Kanamycin resistant gene; mutation; high temperature resistance; coryneform microbe; transformation; gene; ds.
XX		OS	Corynebacterium glutamicum.
XX		XX	Key Location/Qualifiers
XX		XX	1318..2601
XX		XX	/*tag= a
XX		XX	/product= "p48K"
XX		XX	EP1038966-A1.
XX		XX	27-SEP-2000.
XX		XX	16-MAR-2000; 2000EP-00105326.
XX		XX	16-MAR-1999; 99JP-00069896.
XX		XX	(AJIN) AJINOMOTO CO INC.
XX		XX	Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;
XX		XX	WPI; 2000-573832/54.
XX		XX	P-PSDB; AAY97536.
XX		XX	Plasmids containing a temperature sensitive replication control regions useful for breeding microorganisms for the production of amino acids by fermentation.
XX		XX	Claim 7; Page 14-17; 29pp; English.
XX		XX	This sequence encodes the Brevibacterium lactofermentum p48K protein. The invention relates to a plasmid containing a temperature sensitive replication control region (TSRCR) and a marker gene (MG). The TSRCR is derived from plasmid pM330 harboured by Brevibacterium lactofermentum (ATCC 13869) and allows the plasmid to replicate autonomously at a low temperature but does not allow the plasmid to replicate autonomously at an elevated temperature in coryneform bacteria within a temperature range in which the bacteria can grow. The plasmid can be used for modifying a chromosomal gene in a coryneform bacterium, which may be used for the production of useful substances, such as amino acids, by fermentation, to change their genetic traits. Therefore, the plasmid can be used for breeding microorganisms for the production of amino acids by fermentation. The plasmid comprises a TSRCR that allows the plasmid to replicate autonomously at an elevated temperature in coryneform bacteria within a temperature range in which the bacteria can grow. (Updated on 15 -SEP-2003 to standardise OS field)
XX		XX	Sequence 4447 BP; 824 A; 997 C; 1378 G; 1248 T; 0 U; 0 Other;
XX		XX	Query Match 3.3%; Score 37.4; DB 3; Length 4447;
XX		XX	Best Local Similarity 53.0%; Pred. No. 1.7; Indels 0; Gaps 0;
XX		XX	Matches 80; Conservative 0; Mismatches 71;
XX		XX	QY 493 AACGGCTGGACGCTCCACGGTTCACGGCTACTCATGTTTCAGTGGTACGTGAGTACGAGAC 552
XX		XX	1795 AACGGTGGACGCTTCACGGTTCACGGTTCAGTGGTTCAGTGGTACGTGAGTACGAGAT 1854
XX		XX	QY 553 ATCCTCGAATCTTCTCGATGCGATGTTTCGATCGGTGGACTTCCAACTCGTATCTCTG 612
XX		XX	1855 GAACCTCAAGGCTTTGAGATTCAGTGTTCCTCCGCTGGTCTGCTGGTGTGTTAAGGCC 1914
XX		XX	QY 613 GGATTTGCTGGCCACTACGTAATTCGGGTG 643
XX		XX	1915 GGTATGGACGGCCACTCGGTGAGCAGCGGG 1945
XX		XX	RESULT 14
XX		XX	ABL49734
XX		XX	ID ABL49734 standard; DNA; 4447 BP.
XX		XX	AC ABL49734;
XX		XX	XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 23:26:20 ; Search time 119 Seconds
(without alignments)
6809.238 Million cell updates/sec

Title: US-10-007-527A-1

Perfect score: 1140

Sequence: 1 atgaccagcgaagtgtgta.....taatatcggttcgaagtga 1140

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37.4	3.3	4447	3	US-09-521-668B-17
2	37.4	3.3	4447	3	US-09-521-668B-17
3	34.8	3.1	1631	4	US-09-492-709A-208
4	34.6	3.0	1505	1	US-07-915-246-1
5	34.4	3.0	7218	1	US-08-232-463-14
6	33.6	2.9	1134	3	US-09-482-180A-3
7	33.2	2.9	1780	3	US-08-813-150-3
8	33.2	2.9	1780	4	US-09-546-553-3
9	33.2	2.9	1844	4	US-10-140-002-475
10	33.2	2.9	1867	4	US-09-482-273-81
11	32.8	2.9	1987	1	US-08-722-001-26
12	32.8	2.9	2793	1	US-08-209-747-1
13	32.8	2.9	2793	1	US-08-458-298-1
14	32.6	2.9	364	4	US-09-621-912-49
15	32.6	2.9	1431	4	US-09-614-912-49
16	32.6	2.9	1436	4	US-09-614-912-35
17	32.4	2.8	430	4	US-09-621-976-1656
18	32.2	2.8	364	4	US-09-513-999C-30664
19	32.2	2.8	1182	4	US-09-602-787A-663
20	31.8	2.8	909	4	US-09-270-767-30371
21	31.8	2.8	1724	4	US-09-270-767-14243
22	31.8	2.8	1944	4	US-09-489-039A-1217
23	31.6	2.8	505	4	US-09-621-976-15639
24	31.6	2.8	1356	4	US-09-252-991A-14918
25	31.6	2.8	1910	4	US-09-270-767-13531
26	31.4	2.8	4403765	3	US-09-103-840A-2
27	31.4	2.8	4411529	3	US-09-103-840A-1

28	31.2	2.7	767	4	US-09-270-767-12088	Sequence 12088, A
29	31.2	2.7	1599	4	US-09-252-991A-13921	Sequence 13921, A
C 30	31.2	2.7	2280	4	US-09-252-991A-13535	Sequence 13535, A
31	31.2	2.7	5741	1	US-07-706-699-4	Sequence 4, Appli
32	31.2	2.7	5741	1	US-07-998-931-4	Sequence 4, Appli
33	31	2.7	409	4	US-09-270-767-29523	Sequence 29523, A
C 34	31	2.7	3401	4	US-09-907-794A-249	Sequence 249, App
C 35	31	2.7	3401	4	US-09-905-125A-249	Sequence 249, App
C 36	31	2.7	3401	4	US-09-902-775A-249	Sequence 249, App
C 37	31	2.7	3401	4	US-09-906-700-249	Sequence 249, App
C 38	31	2.7	3401	4	US-09-903-603A-249	Sequence 249, App
39	30.8	2.7	1686	4	US-09-489-039A-1299	Sequence 1299, App
40	30.8	2.7	3651	4	US-09-614-221A-408	Sequence 408, App
41	30.8	2.7	5820	3	US-09-029-213B-7	Sequence 7, Appli
C 42	30.6	2.7	399	4	US-09-621-976-8976	Sequence 8976, App
C 43	30.6	2.7	1708	5	PCT-US95-02481-3	Sequence 3, Appli
C 44	30.6	2.7	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 45	30.6	2.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-521-668B-17
; Sequence 17, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAHIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521,668B
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: JP 11-69896
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 4447
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1318)..(2598)
US-09-521-668B-17

Query Match 3.3%; Score 37.4; DB 3; Length 4447;
Best Local Similarity 53.0%; Pred. No. 0.24;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 493 AACGGTGGACGTCACGTTACGGCGTACTCATGTCAGTGGTGGACGTGAGAGAAC 552
DB 1795 AACGGTGGACGTTGACCGCAACATGCTGTTCTTGGATCGTCCACTGTCTGACGAT 1854
QY 553 ATCTCGAATCCTTCTCGGATCGATGTTGCGATCGGTGGACTTCCAACTCGTATCTCTG 612
DB 1855 GAATCAAGCGGTTGAGGATTCATGTTTCCCGTGGTCTGCTGGTGTAAAGGCC 1914
QY 613 GGATTTGCTGCGCCACTACGTAATTCGGGTG 643
DB 1915 GGATGAGCGCGCCACTGCGTGAGCAGCGGG 1945

RESULT 2
US-09-521-668B-19
; Sequence 19, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN

	Query Match	3.1%	Score 34.8	DB 4	Length 1631
	Best Local Similarity	50.0%	Pred. No. 0.94		
	Matches 87	Conservative	0	Mismatches 87	Indels 0
	Gaps	0			
Qy	219	CGGAAAGCGGTGGATCTGCCCTCTGTGCGGGAAAAGTCGGTGCACATCGTCGACGACA	278		
Db	1184	CGGAAATGTCCTGGAGCGACCGGAGGTGCAAAAAAATGGCGCGAAACGTGATCTGCCGC	1243		
	279	AATTTCACAGTCTGTGCTCATCAACTCGGACCTGGATCTCTTGCAGTGGTGACGATGAC	338		

	Query Match	3.0%;	Score 34.6;	DB 1;	Length 1505;
	Best Local Similarity	58.1%;	Pred. No. 1;		
	Matches 61;	Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0;
Qy	706	ATTGCATCTCGCGTGTGATGGAGGTTGGTAGTGGCGACGGAAAAAGTGTGCGACATGGC	765		
Db	699	ATTGGAGTTGATGTCGTGTGTTGTTTGGAGGTGTGTCGCGAGAGAGGTGGTCTGGT	758		
Qy	766	AACCGTGCACCCCTGGGAAATCCGCTGTTTGATGTCAGTGGCGCGGGGAT	810		
Db	759	TACGCTGACCTGAAGCATTGGTTGGAGGAGGAGCGCGTGGACAT	803		

RESULT 5

ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/813,150
FILING DATE: 07-MAR-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0693
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1780 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 66..1766
US-08-813-150-3

Query Match 2.9%; Score 33.2; DB 3; Length 1780;
Best Local Similarity 55.1%; Pred. No. 3.2;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 640 GGTGCTCTCGATGTACGAAGAATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 699
DB 1016 GGTGCTCTCGATGTACGAAGAATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 1075
QY 700 ACGAAATTTGATCTGGCGTGTGATGAGGTTGGTAGGCGGACGAGAAAGTGGTC 757
DB 1076 CGCCACATGACGAGGAGCGCTGCGGAGGCTGCACACTACGTGCGCGCCACCAAGTGTTC 1133

RESULT 8
US-09-546-553-3
Sequence 3, Application US/09546553
Patent No. 6518405
GENERAL INFORMATION:
APPLICANT: Mueller, Christopher
APPLICANT: Lebecque, Serge J.E.
APPLICANT: Liu, Yong-Jun
APPLICANT: Dowling, Lynette M.
APPLICANT: Huffine, Constance M.
APPLICANT: Gorman, Daniel M.
TITLE OF INVENTION: MAMMALIAN PROTEINASES; OXIDOREDUCTASES;
TITLE OF INVENTION: RELATED REAGENTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/546,553
FILING DATE: 10-APR-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/813,150
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0693
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1780 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 66..1766
US-09-546-553-3

Query Match 2.9%; Score 33.2; DB 4; Length 1780;
Best Local Similarity 55.1%; Pred. No. 3.2;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 640 GGTGCTCTCGATGTACGAAGAATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 699
DB 1016 GGTGCTCTCGATGTACGAAGAATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 1075
QY 700 ACGAAATTTGATCTGGCGTGTGATGAGGTTGGTAGGCGGACGAGAAAGTGGTC 757
DB 1076 CGCCACATGACGAGGAGCGCTGCGGAGGCTGCACACTACGTGCGCGCCACCAAGTGTTC 1133

RESULT 9
US-10-140-002-475
Sequence 475, Application US/10140002
Patent No. 6725730
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC59
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 475
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-475

Db 209 RKSSKKRYATYYSCMMRWKYYCMMRSATYSGCMMRWYCYSCMMRSYCTSYSGKCS 150
QY 858 ACCTCGCGGCAATCCGCTGGTCCCGTGGATTGCGTCCCGAGCTGGTCTTTGGGGCAGAACT 917
Db 149 CTGWGKCKCYRMYRGRWYKYSRGARRYTKSRGWCRCSTKRYKTKASWGAAGTQWC 90
QY 918 AACAGATGCTCAGATCGTTGAGCAGGAAGAATCTGCCCGGTCATGTTGGATCATCTCC 977
Db 89 CMRWGSTGASVMRWKYSYKRWKWSWKMKGYSWYSWMSGSCWCKSCGSSSYCRSYSS 30
QY 978 GGCGCGATCGTGATGATCGGACTT 1006
Db 29 MGKSRCTCCMGCGGCTTCMKSCRYT 1

RESULT 15

US-09-614-912-49/c
; Sequence 49, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cai, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 49
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1309)
; NAME/KEY: unsure
; LOCATION: (1339)
; NAME/KEY: unsure
; LOCATION: (1351)
; NAME/KEY: unsure
; LOCATION: (1402)
; NAME/KEY: unsure
; LOCATION: (1429)
US-09-614-912-49

Query Match 2.9%; Score 32.6; DB 4; Length 1431;
Best Local Similarity 49.7%; Pred. No. 4.5;
Matches 83; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 890 GTGCCCGAGCTGCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAT 949
Db 850 GCGCCCGCGCGGCGCTGATGAGGCGCACCGCAGCGTGGCGCGTGCACACCTGGAGT 791

QY 950 CTGCCCGGTCATGGTTGCGATCAATCCGGCGCATCGTTGATGATGATTCGGAATTGTG 1009
Db 790 CCGCCACGTCGTCCTGTCAGGATGGTGAGCGACGTTGGGGTCCCAATCGGCCCGCTG 731
QY 1010 CGCCTTACGTCTTCGGCGAGATCCTCGGACTCTCGTCTGAAAGCTGGGGCG 1056
Db 730 CCCAGCGTTCGTTAGGCGCGCTGGCACGCGCGGTAGTAGTTGAGGCG 684

Search completed: October 23, 2004, 02:20:26
Job time : 129 secs


```
QY 121 GAAACATTAAACGCTGTGGCGCGGATTTCTGGGTGAACGGTGTGACCAATTGTCAAC 180
DB |||||
QY 121 GAAACATTAAACGCTGTGGCGCGGATTTCTGGGTGAACGGTGTGACCAATTGTCAAC 180
DB |||||
QY 181 GGTCCGAAAGTTCTGGATTTCGGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 240
DB |||||
QY 181 GGTCCGAAAGTTCTGGATTTCGGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 240
DB |||||
QY 241 TGTCTGGCGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 241 TGTCTGGCGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 301 CAACTCGGAGCTGATCTGTCGGATGTGACGATGACCATCGGCATACACGCTGTGAC 360
DB |||||
QY 301 CAACTCGGAGCTGATCTGTCGGATGTGACGATGACCATCGGCATACACGCTGTGAC 360
DB |||||
QY 361 CGGCTCCAGCACTATGGAATGGAATTTTCGGAGGCTTCGAAAGCTGCGAACCGTCT 420
DB |||||
QY 361 CGGCTCCAGCACTATGGAATGGAATTTTCGGAGGCTTCGAAAGCTGCGAACCGTCT 420
DB |||||
QY 421 CGTTCGGGTACGGAAGTGAATGTACGGCTCGGACGATACGTCGGCGCTGTGAATTC 480
DB |||||
QY 421 CGTTCGGGTACGGAAGTGAATGTACGGCTCGGACGATACGTCGGCGCTGTGAATTC 480
DB |||||
QY 481 ACTCAGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 481 ACTCAGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 541 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGATGTTCTCGATCGGCTGTCCAAA 600
DB |||||
QY 541 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGATGTTCTCGATCGGCTGTCCAAA 600
DB |||||
QY 601 CTCGATCTCTGGGATTTCTGGGATTCGAGTACGTCGGCGCTGTGAATTC 720
DB |||||
QY 601 CTCGATCTCTGGGATTTCTGGGATTCGAGTACGTCGGCGCTGTGAATTC 720
DB |||||
QY 661 ATCGCGGTGAGAGTGTGATCAAGTTCTCGCTGGTATCTGACGAAATTTGCATCTGCG 720
DB |||||
QY 661 ATCGCGGTGAGAGTGTGATCAAGTTCTCGCTGGTATCTGACGAAATTTGCATCTGCG 720
DB |||||
QY 721 GGTATGGAGTTGGTAGTGGGACGAAAGTGGTTCGATGCGAACCGTGCACCCCTGG 780
DB |||||
QY 721 GGTATGGAGTTGGTAGTGGGACGAAAGTGGTTCGATGCGAACCGTGCACCCCTGG 780
DB |||||
QY 781 GAAATCGCTGTGATGAGTGGGACGAAAGTGGTTCGATGCGAACCGTGCACCCCTGG 840
DB |||||
QY 781 GAAATCGCTGTGATGAGTGGGACGAAAGTGGTTCGATGCGAACCGTGCACCCCTGG 840
DB |||||
QY 841 GAGTTTGGTTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGGATTCGCTGCGAGCT 900
DB |||||
QY 841 GAGTTTGGTTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGGATTCGCTGCGAGCT 900
DB |||||
QY 901 GGTCTTGGGGCAGAACTAACAGATGCTCAGATGCTGTTGAGCAGGAAATTCGCCCGGTC 960
DB |||||
QY 901 GGTCTTGGGGCAGAACTAACAGATGCTCAGATGCTGTTGAGCAGGAAATTCGCCCGGTC 960
DB |||||
QY 961 ATGGTTGCGATCAATTCGGCGGATCGTGATGATGATTCGGAATTCGCGCTTACGTC 1020
DB |||||
QY 961 ATGGTTGCGATCAATTCGGCGGATCGTGATGATGATTCGGAATTCGCGCTTACGTC 1020
DB |||||
QY 1021 TTCCGCGAGATCTCCGACTCGTTCGAAGCTGGCGGACCTTGGGAAATCTTCGTGATCAC 1080
DB |||||
QY 1021 TTCCGCGAGATCTCCGACTCGTTCGAAGCTGGCGGACCTTGGGAAATCTTCGTGATCAC 1080
DB |||||
QY 1081 TTGCATATTCGATTCGCGGACGAGATTCGCGGCCCCCGGATATTCGTTTCGCAAGTGA 1140
DB |||||
QY 1081 TTGCATATTCGATTCGCGGACGAGATTCGCGGCCCCCGGATATTCGTTTCGCAAGTGA 1140
DB |||||
```

RESULT 2

US-10-007-452-1

; Sequence 1, Application US/10007452

; Publication No. US20030093701A1

```
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; PRIORITY FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-007-452-1
```

```
Query Match 100.0%; Score 1140; DB 14; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCAAGCTAAGTGTGAAACACCTTTCCGGCAAGACCCGGCTTCCCGTCTCGTGTGCG 60
DB |||||
QY 1 ATGACCAAGCTAAGTGTGAAACACCTTTCCGGCAAGACCCGGCTTCCCGTCTCGTGTGCG 60
DB |||||
QY 61 TCCGATAAGCGCGGATCCGGACGAACTGCGACCCAACTTCAACAAATCAACACGTCAC 120
DB |||||
QY 61 TCCGATAAGCGCGGATCCGGACGAACTGCGACCCAACTTCAACAAATCAACACGTCAC 120
DB |||||
QY 121 GAAACATTAAACGCTGTGGCGCGGATTTCTGGGTGAACGGTGTGACCAATTGTCAAC 180
DB |||||
QY 121 GAAACATTAAACGCTGTGGCGCGGATTTCTGGGTGAACGGTGTGACCAATTGTCAAC 180
DB |||||
QY 181 GGTCCGAAAGTTCGGAATTCGAGGCTTCTGCTCGGAAAGGCTGGATCTGCCCC 240
DB |||||
QY 181 GGTCCGAAAGTTCGGAATTCGAGGCTTCTGCTCGGAAAGGCTGGATCTGCCCC 240
DB |||||
QY 241 TGTCTGGCGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 241 TGTCTGGCGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 301 CAACTCGGAGCTGATCTGTCGGATGTGACGATGACCATCGGCATACACGCTGTGAC 360
DB |||||
QY 301 CAACTCGGAGCTGATCTGTCGGATGTGACGATGACCATCGGCATACACGCTGTGAC 360
DB |||||
QY 361 CGGCTCCAGCACTATGGAATGGAATTTTCGGAGGCTTCGAAAGCTGCGAACCGTCT 420
DB |||||
QY 361 CGGCTCCAGCACTATGGAATGGAATTTTCGGAGGCTTCGAAAGCTGCGAACCGTCT 420
DB |||||
QY 421 CGTTCGGGTACGGAAGTGAATGTACGGCTCGGACGATACGTCGGCGCTGTGAATTC 480
DB |||||
QY 421 CGTTCGGGTACGGAAGTGAATGTACGGCTCGGACGATACGTCGGCGCTGTGAATTC 480
DB |||||
QY 481 ACTCAGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 481 ACTCAGGAAAGTGGTGACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||
QY 541 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGATGTTCTCGATCGGCTGTCCAAA 600
DB |||||
QY 541 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGATGTTCTCGATCGGCTGTCCAAA 600
DB |||||
QY 601 CTCGATCTCTGGGATTTCTGGGATTCGAGTACGTCGGCGCTGTGAATTC 720
DB |||||
QY 601 CTCGATCTCTGGGATTTCTGGGATTCGAGTACGTCGGCGCTGTGAATTC 720
DB |||||
QY 661 ATCGCGGTGAGAGTGTGATCAAGTTCTCGCTGGTATCTGACGAAATTTGCATCTGCG 720
DB |||||
QY 661 ATCGCGGTGAGAGTGTGATCAAGTTCTCGCTGGTATCTGACGAAATTTGCATCTGCG 720
DB |||||
QY 721 GGTATGGAGTTGGTAGTGGGACGAAAGTGGTTCGATGCGAACCGTGCACCCCTGG 780
DB |||||
```

```
Db 721 GGTATGGAGTTGGTAGTGGCGACGGAAGAGTGGTGCAGATGCCAACCGTGCACCGTGG 780
Qy 781 GAAATCGCTGTGATCGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Db 781 GAAATCGCTGTGATCGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Qy 841 GAGTTTGGTTCCGATGGGAGTGGGCAATCGGTGTGTCCTCGTGGATTTGCGTCCCGAGCT 900
Db 841 GAGTTTGGTTCCGATGGGAGTGGGCAATCGGTGTGTCCTCGTGGATTTGCGTCCCGAGCT 900
Qy 901 GGTCTGGGGCAGAACTAACAGATCTCAGATCGTTGAGCAGAGAAATCTGCCCGGTC 960
Db 901 GGTCTGGGGCAGAACTAACAGATCTCAGATCGTTGAGCAGAGAAATCTGCCCGGTC 960
Qy 961 ATGGTGGGATCATTCGCGGCGGATCGGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Db 961 ATGGTGGGATCATTCGCGGCGGATCGGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Qy 1021 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTGATCAC 1080
Db 1021 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTGATCAC 1080
Qy 1081 TTGCATTATCGATTGCCCGCAGCGGATGTGGCGCCCGGATAATATCGTTTCGCAAGTGA 1140
Db 1081 TTGCATTATCGATTGCCCGCAGCGGATGTGGCGCCCGGATAATATCGTTTCGCAAGTGA 1140

RESULT 3
US-10-415-562A-1
; Sequence 1, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: Cl1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; PRIOR FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-415-562A-1

Query Match
Best Local Similarity 100.0%; Score 1140; DB 17; Length 1140;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACACGGTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
Db 1 ATGACACGGTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
Qy 61 TCCGATAGCGCGGCATCCGCGACGAACTGGACCCAACTTCAACAAATCACCACGTCA 120
Db 61 TCCGATAGCGCGGCATCCGCGACGAACTGGACCCAACTTCAACAAATCACCACGTCA 120
Qy 121 GAAACATTTAAACGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCATTTGCAAC 180
Db 121 GAAACATTTAAACGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCATTTGCAAC 180
Qy 181 GGTCCGAAAGTTCTGATTCGAGGCTTCTGTTCCGCGAAAGGCTGGATCTGCCCC 240
Db 181 GGTCCGAAAGTTCTGATTCGAGGCTTCTGTTCCGCGAAAGGCTGGATCTGCCCC 240
Qy 241 TGCTGTCGGGAAAGTCTGGTGCACATCTGCGTGAACGAAATTTCTCAAGTTGTGCTCAT 300
Db 241 TGCTGTCGGGAAAGTCTGGTGCACATCTGCGTGAACGAAATTTCTCAAGTTGTGCTCAT 300
Qy 301 CAACTCGGAGTCTGATCTGTTGGATGGTACCATGACCATGCGCCATACAGTGTGTCAG 360
Db 301 CAACTCGGAGTCTGATCTGTTGGATGGTACCATGACCATGCGCCATACAGTGTGTCAG 360
```

```
Db 301 CAACTCGGAGTCTGATCTGTTGGATGGTACCATGACCATGCGCCATACAGTGTGTCAG 360
Qy 361 CGGCTCCACGACTATGAGTGGACTTTTCGGAGCGCTGGAAAGCTCGGACCAACGGTCTGT 420
Db 361 CGGCTCCACGACTATGAGTGGACTTTTCGGAGCGCTGGAAAGCTCGGACCAACGGTCTGT 420
Qy 421 CTTTGGCGTACGAAACGTGAAATGTACGGCTCGGACGATACGTGGCGCTGTGAAATC 480
Db 421 CTTTGGCGTACGAAACGTGAAATGTACGGCTCGGACGATACGTGGCGCTGTGAAATC 480
Qy 481 ACTCAGGAAAGAAACGGCTGGCACTCCACGTTTACGGCTCGGACGATACGTGGCTGAC 540
Db 481 ACTCAGGAAAGAAACGGCTGGCACTCCACGTTTACGGCTCGGACGATACGTGGCTGAC 540
Qy 541 GTGAGTGAGAAACATCTCGAATCTTCTCGAATCTTCTCGAATCTTCTCGAATCTTCTCGAAT 600
Db 541 GTGAGTGAGAAACATCTCGAATCTTCTCGAATCTTCTCGAATCTTCTCGAATCTTCTCGAAT 600
Qy 601 CTCTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGGTGCTCGATGTACGAAAG 660
Db 601 CTCTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGGTGCTCGATGTACGAAAG 660
Qy 661 ATCGCGGTGAAGCTGATCAAGTTCTCGTGCCTATCTGACGAAATTTGCACTGCGGCTT 720
Db 661 ATCGCGGTGAAGCTGATCAAGTTCTCGTGCCTATCTGACGAAATTTGCACTGCGGCTT 720
Qy 721 GGTATGGAGGTTGGTAGTGGCGACGGAAGAGTGGTGCAGATGGCAACCGTGCACCGTGG 780
Db 721 GGTATGGAGGTTGGTAGTGGCGACGGAAGAGTGGTGCAGATGGCAACCGTGCACCGTGG 780
Qy 781 GAAATCGCTGTGATGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Db 781 GAAATCGCTGTGATGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Qy 841 GAGTTTGGTTCCGATGGGAGTGGCGCAATCGGTGTGTCCTCGTGGATTTGCGTCCCGAGCT 900
Db 841 GAGTTTGGTTCCGATGGGAGTGGCGCAATCGGTGTGTCCTCGTGGATTTGCGTCCCGAGCT 900
Qy 901 GGTCTTGGCGGACAACTAACAGATCTCAGATGATTCGGACTTGTGGCGCTTACGTC 1020
Db 901 GGTCTTGGCGGACAACTAACAGATCTCAGATGATTCGGACTTGTGGCGCTTACGTC 1020
Qy 961 ATGGTGGGATCATTCGCGGCGGATCGGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 1080
Db 961 ATGGTGGGATCATTCGCGGCGGATCGGTGGATGATGATTCGGACTTGTGGCGCTTACGTC 1080
Qy 1021 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTGATCAC 1140
Db 1021 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTGATCAC 1140
Qy 1081 TTGCATTATCGATTGCCCGCAGCGGATGTGGCGCCCGGATAATATCGTTTCGCAAGTGA 1140
Db 1081 TTGCATTATCGATTGCCCGCAGCGGATGTGGCGCCCGGATAATATCGTTTCGCAAGTGA 1140
```

RESULT 4

```
US-10-007-527A-5/c
; Sequence 5, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: Cl1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
```

```

; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-007-527A-5

Query Match      100.0%; Score 1140; DB 14; Length 6334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGAGTAAGTGTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCTGTCG 60
DB 3051 ATGACCGAGTAAGTGTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCTGTCG 2992

QY 61 TCCGATAGCGGATCGGACGACGAACTGCGACCAAACTTCAACAAATCACACGTC 120
DB 2991 TCCGATAGCGGATCGGACGACGAACTGCGACCAAACTTCAACAAATCACACGTC 2932

QY 121 GAAACATTTAAAGCGCTGTGCGCGGCGGATTTCTGGCGTGAACGGTGTGCAAC 180
DB 2931 GAAACATTTAAAGCGCTGTGCGCGGCGGATTTCTGGCGTGAACGGTGTGCAAC 2872

QY 181 GGTCCGAAAGTTCTGGATTTCGGAGCGCTTCCTCGGAAAGGGCTGGATCTGCC 240
DB 2871 GGTCCGAAAGTTCTGGATTTCGGAGCGCTTCCTCGGAAAGGGCTGGATCTGCC 2812

QY 241 TCGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 300
DB 2811 TCGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 2752

QY 301 CAACTCGGACCTGGATCTGTCGATGCGATGCGATGCGATGCGATGCGATGCGAT 360
DB 2751 CAACTCGGACCTGGATCTGTCGATGCGATGCGATGCGATGCGATGCGATGCG 2692

QY 361 CGGCTCCACGACCTATGGAATTCGGACCTTCGGACGCTTCGGACGCTTCGGAC 420
DB 2691 CGGCTCCACGACCTATGGAATTCGGACGCTTCGGACGCTTCGGACGCTTCGG 2632

QY 421 CGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 300
DB 2811 CGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 2752

QY 301 CAACTCGGACCTGGATCTGTCGATGCGATGCGATGCGATGCGATGCGATGCG 360
DB 2751 CAACTCGGACCTGGATCTGTCGATGCGATGCGATGCGATGCGATGCGATGCG 2692

QY 361 CGGCTCCACGACCTATGGAATTCGGACCTTCGGACGCTTCGGACGCTTCGGAC 420
DB 2691 CGGCTCCACGACCTATGGAATTCGGACGCTTCGGACGCTTCGGACGCTTCGG 2632

QY 421 GGTTCGGCTACGGAACGTGAATGTACGGCTGCGACGATACGTGCGGCTGTTGA 480
DB 2631 GGTTCGGCTACGGAACGTGAATGTACGGCTGCGACGATACGTGCGGCTGTTGA 2572

QY 481 ACTCAGGAAAAAAGCGGTGGCACTGTCAGTTCAGCGCTTCACTGTCAGTGGT 540
DB 2571 ACTCAGGAAAAAAGCGGTGGCACTGTCAGTTCAGCGCTTCACTGTCAGTGGT 2512

QY 541 GTGAGTGAGACATCTCGAATCTTCTCGATGCGATGCGATGCGATGCGATGCG 600
DB 2511 GTGAGTGAGACATCTCGAATCTTCTCGATGCGATGCGATGCGATGCGATGCG 2452

QY 601 CTCGTATCTCTGGATTTCTCGGCACTTACGTAATTCGGGTGGTCTCGATGCA 560
DB 2451 CTCGTATCTCTGGATTTCTCGGCACTTACGTAATTCGGGTGGTCTCGATGCA 2392

QY 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCATCTGGCG 720
DB 2391 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCATCTGGCG 2332

QY 721 GGTATGAGGTTGGTAGTGGCAGGAAAAAGTGGTGGATGCGCAACCGTGCACCT 780
DB 2331 GGTATGAGGTTGGTAGTGGCAGGAAAAAGTGGTGGATGCGCAACCGTGCACCT 2272

QY 781 GAAATCGCTTTGATGAGTGGCGGGGATTCACAAAGCGTTGGAATCTGGCGAGA 840
DB 2271 GAAATCGCTTTGATGAGTGGCGGGGATTCACAAAGCGTTGGAATCTGGCGAGA 2212

QY 841 GAGTTGGTTGATGAGGAGCTCGGCAATCGCTGGTTCGCTGGATTGCGTGGCGAG 900
DB 2211 GAGTTGGTTGATGAGGAGCTCGGCAATCGCTGGTTCGCTGGATTGCGTGGCGAG 2152

QY 901 GGTCTCGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCC 960
DB 2151 GGTCTCGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCC 2092

```

```

RESULT 5
US-10-007-452-5/c
; Sequence 5, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francoise
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-007-452-5

```

```

Query Match      100.0%; Score 1140; DB 14; Length 6334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGAGTAAGTGTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCTGTCG 60
DB 3051 ATGACCGAGTAAGTGTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCTGTCG 2992

QY 61 TCCGATAGCGGATCGGACGACGAACTGCGACCAAACTTCAACAAATCACACGTC 120
DB 2991 TCCGATAGCGGATCGGACGACGAACTGCGACCAAACTTCAACAAATCACACGTC 2932

QY 121 GAAACATTTAAAGCGCTGTGCGCGGCGGATTTCTGGCGTGAACGGTGTGCAAC 180
DB 2931 GAAACATTTAAAGCGCTGTGCGCGGCGGATTTCTGGCGTGAACGGTGTGCAAC 2872

QY 181 GGTCCGAAAGTTCTGGATTTCGGAGCGCTTCCTCGGAAAGGGCTGGATCTGCC 240
DB 2871 GGTCCGAAAGTTCTGGATTTCGGAGCGCTTCCTCGGAAAGGGCTGGATCTGCC 2812

QY 241 TCGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 300
DB 2811 TCGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 2752

QY 301 CAACTCGGACCTGGATCTGTCGATGCGATGCGATGCGATGCGATGCGATGCGAT 360
DB 2751 CAACTCGGACCTGGATCTGTCGATGCGATGCGATGCGATGCGATGCGATGCG 2692

QY 361 CGGCTCCACGACCTATGGAATTCGGACCTTCGGACGCTTCGGACGCTTCGGAC 420
DB 2691 CGGCTCCACGACCTATGGAATTCGGACGCTTCGGACGCTTCGGACGCTTCGG 2632

QY 421 CGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 480
DB 2631 CGTGTGCGGAAAAAGTGGTGCACATCGTCGAGCAAAATTTCTCAAGTTGTGCT 2572

QY 481 ACTCAGGAAAAAAGCGGTGGCACTGTCAGTTCAGCGCTTCACTGTCAGTGGT 540

```

```
Db 2571 ACTCAGGAAAAAGCGCTGGCAGCTCCAGCTTCAACGCGTACTCATGTTTCAGTGGTGAC 2512
QY 541 GTGAGTGAGAAATCATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGACTTCCAAA 600
Db 2511 GTGAGTGAGAAATCATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGACTTCCAAA 2452
QY 601 CTCGTATCTCTGGATTTGCTGGCCACTACGTAAATTCGGGTGGTCTCGATGTACGAAG 660
Db 2451 CTCGTATCTCTGGATTTGCTGGCCACTACGTAAATTCGGGTGGTCTCGATGTACGAAG 2392
QY 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCGATCTGGCGTT 720
Db 2391 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCGATCTGGCGTT 2332
QY 721 GGTATGGAGTTGGTAGTGGCGACGAAAGTGTGCGATGGCAACCTGACCTGG 780
Db 2331 GGTATGGAGTTGGTAGTGGCGACGAAAGTGTGCGATGGCAACCTGACCTGG 2272
QY 781 GAAATCGCTGTTGATGCGAGTGGCGGGGATCCACAGCGTTCGAACTGTGGCGAGAAATT 840
Db 2271 GAAATCGCTGTTGATGCGAGTGGCGGGGATCCACAGCGTTCGAACTGTGGCGAGAAATT 2212
QY 841 GAGTTTGGTTCGATGGGACGTTCGGCAATCGCGTGGTCCCGTGGATTCGTCGCCGAGCT 900
Db 2211 GAGTTTGGTTCGATGGGACGTTCGGCAATCGCGTGGTCCCGTGGATTCGTCGCCGAGCT 2152
QY 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATTCGCCCGGTC 960
Db 2151 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATTCGCCCGGTC 2092
QY 961 ATGGTTGGGATCATTCGGCGCGATCGTGGATGATGATTCGGACTTGTGCGCTTACGTC 1020
Db 2091 ATGGTTGGGATCATTCGGCGCGATCGTGGATGATGATTCGGACTTGTGCGCTTACGTC 2032
QY 1021 TTCGCGAGATCCTCGGACTCTCGAAGCTCGGAGCTGGCGGAGTTCGGGAAATTCCTGATCAC 1080
Db 2031 TTCGCGAGATCCTCGGACTCTCGAAGCTCGGAGCTGGCGGAGTTCGGGAAATTCCTGATCAC 1972
QY 1081 TTGCATTTATCGATTCGCCGAGCGATGTGGCGCCCGCGATTAATATCGGTTTCGCAAGTGA 1140
Db 1971 TTGCATTTATCGATTCGCCGAGCGATGTGGCGCCCGCGATTAATATCGGTTTCGCAAGTGA 1912
```

RESULT 6

```
US-10-415-562A-5/c
; Sequence 5, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-415-562A-5
```

```
Query Match 100.0%; Score 1140; DB 17; Length 6334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGGTAAGTGTGAACACTTTCGGCGAAGACCGGCTCCCGTCTCGTGTGCG 60
Db 3051 ATGACCGGTAAGTGTGAACACTTTCGGCGAAGACCGGCTCCCGTCTCGTGTGCG 2992
QY 61 TCCGATAGCGCGGATCCGGCAGCACTGCGACCCAACTTCACAAATCACCACGTCA 120
```

```
Db 2991 TCCGATAGCGCGGATCCGGCAGCACTGCGACCCAACTTCACAAATCACCACGTCA 2932
QY 121 GAAACATTTAAACCCCTGTGCGCGCGGATTTCTTGGGCTGAAACGCTGTGACCATTTGTCAAC 180
Db 2931 GAAACATTTAAACCCCTGTGCGCGCGGATTTCTTGGGCTGAAACGCTGTGACCATTTGTCAAC 2872
QY 181 GGTCCGAAAGGTTCTGGATTCGGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 240
Db 2871 GGTCCGAAAGGTTCTGGATTCGGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 2812
QY 241 TCGTGTGCGGAAAGTTCGGTGCACATCGTGCAGAGAAATTTCTCAAAGTTGTGCTCAT 300
Db 2811 TCGTGTGCGGAAAGTTCGGTGCACATCGTGCAGAGAAATTTCTCAAAGTTGTGCTCAT 2752
QY 301 CAACTCGGACTGGATCTGTTGCGATGTTGACGATGACCATCGCCATACAGCTGGTCAG 360
Db 2751 CAACTCGGACTGGATCTGTTGCGATGTTGACGATGACCATCGCCATACAGCTGGTCAG 2692
QY 361 CGGCTCCAGGACTATGGAATTCGGCAGCTTCGGAAGCTTCGGAACCAACGCTCGT 420
Db 2691 CGGCTCCAGGACTATGGAATTCGGCAGCTTCGGAAGCTTCGGAACCAACGCTCGT 2632
QY 421 CGTTCGGCTACGGAACGTGAAATGTACGCTCGCAGCGATACGTCGGCGCTGTGAAATC 480
Db 2631 CGTTCGGCTACGGAACGTGAAATGTACGCTCGCAGCGATACGTCGGCGCTGTGAAATC 2572
QY 481 ACTCAGGAAAGGCTGGCAGCTTCACGCTTCACGCGCTACTCATGTTTCAGTGGTGAC 540
Db 2571 ACTCAGGAAAGGCTGGCAGCTTCACGCTTCACGCGCTACTCATGTTTCAGTGGTGAC 2512
QY 541 GTGAGTGAGAAATCCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGACTTCCAAA 600
Db 2511 GTGAGTGAGAAATCCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGACTTCCAAA 2452
QY 601 CTCGTATCTCTGGGATTTGCTGGCGCACTACGTAATTCGGGTGGTTCGATGTACGAAG 660
Db 2451 CTCGTATCTCTGGGATTTGCTGGCGCACTACGTAATTCGGGTGGTTCGATGTACGAAG 2392
QY 661 ATCGCGGTGAGCTGATCAAGTTCTCGCTCGTATCTGACGAAATTCGATCTCGCGTT 720
Db 2391 ATCGCGGTGAGCTGATCAAGTTCTCGCTCGTATCTGACGAAATTCGATCTCGCGTT 2332
QY 721 GGTATGAGGTTGTTAGTGGCGACGGAAGGTTGTCATGTCGACATGCGACCCCTGG 780
Db 2331 GGTATGAGGTTGTTAGTGGCGACGGAAGGTTGTCATGTCGACATGCGACCCCTGG 2272
QY 781 GAAATCGCTGTTGATGCGAGTGGCGGGATCCACAGCGTTCGAACTGTGGCGAGAAATT 840
Db 2271 GAAATCGCTGTTGATGCGAGTGGCGGGATCCACAGCGTTCGAACTGTGGCGAGAAATT 2212
QY 841 GAGTTTGGTTCGATGGGACGTTCGGCAATCGCGTGGTCCCGTGGATTCGTCGCCGAGCT 900
Db 2211 GAGTTTGGTTCGATGGGACGTTCGGCAATCGCGTGGTCCCGTGGATTCGTCGCCGAGCT 2152
QY 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATTCGCCCGGTC 960
Db 2151 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATTCGCCCGGTC 2092
QY 961 ATGGTTGCGATCATTCGGCGCGATCGTGGATGATGATTCGGAATTCGTCGCCCTTACGTC 1020
Db 2091 ATGGTTGCGATCATTCGGCGCGATCGTGGATGATGATTCGGAATTCGTCGCCCTTACGTC 2032
QY 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATTCCTGATGATCAC 1080
Db 2031 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATTCCTGATGATCAC 1972
QY 1081 TTGCATTTATCGATTCGCCGAGCGATGTGGCGCCCGCGATTAATATCGGTTTCGCAAGTGA 1140
Db 1971 TTGCATTTATCGATTCGCCGAGCGATGTGGCGCCCGCGATTAATATCGGTTTCGCAAGTGA 1912
```

RESULT 7

```
US-10-007-527A-7/c
; Sequence 7, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-527A-7

Query Match      100.0%; Score 1140; DB 14; Length 9652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGACCAGCGTAAAGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
DB      5052 ATGACCAGCGTAAAGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 4993

QY      61  TCGGATAAGCGCGGCATCCGGCAGCACTGCGACCCCAAACTTCAACAAATCACCAGTCA 120
DB      4992 TCGGATAAGCGCGGCATCCGGCAGCACTGCGACCCCAAACTTCAACAAATCACCAGTCA 4933

QY      121 GAAACATTTAAACCGCTGTGGCGCGGATTTCTGGCGTGAAACCGTGTGACCATTTGCAAC 180
DB      4932 GAAACATTTAAACCGCTGTGGCGCGGATTTCTGGCGTGAAACCGTGTGACCATTTGCAAC 4873

QY      181 GGTCCGGAAGGTTCTGGATTTCGGAGGCTTCTGTTCCCTCGGAAAGGGCTGGATCTGGCCC 240
DB      4872 GGTCCGGAAGGTTCTGGATTTCGGAGGCTTCTGTTCCCTCGGAAAGGGCTGGATCTGGCCC 4813

RESULT 8
US-10-007-452-7/c
; Sequence 7, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-452-7

Query Match      100.0%; Score 1140; DB 14; Length 9652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGACCAGCGTAAAGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
DB      5052 ATGACCAGCGTAAAGTCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 4993

QY      61  TCGGATAAGCGCGGCATCCGGCAGCACTGCGACCCCAAACTTCAACAAATCACCAGTCA 120
DB      4992 TCGGATAAGCGCGGCATCCGGCAGCACTGCGACCCCAAACTTCAACAAATCACCAGTCA 4933

QY      121 GAAACATTTAAACCGCTGTGGCGCGGATTTCTGGCGTGAAACCGTGTGACCATTTGCAAC 180
DB      4932 GAAACATTTAAACCGCTGTGGCGCGGATTTCTGGCGTGAAACCGTGTGACCATTTGCAAC 4873

QY      181 GGTCCGGAAGGTTCTGGATTTCGGAGGCTTCTGTTCCCTCGGAAAGGGCTGGATCTGGCCC 240
DB      4872 GGTCCGGAAGGTTCTGGATTTCGGAGGCTTCTGTTCCCTCGGAAAGGGCTGGATCTGGCCC 4813
```


QY 241 TGCTGCGGAAAGTCGGTGCACATCGTGCAGCAAGAAATTTCTCAAGTTGTGCTCAT 300
DB |||||||
QY 4812 TGCTGTGCGGAAAGTCGGTGCACATCGTGCAGCAAGAAATTTCTCAAGTTGTGCTCAT 4753
DB |||||||
QY 301 CAACTCGGGAGTGGATCTGTGGCATGGTGCAGTACCAATCGCCCATACAGCTGGTCAG 360
DB |||||||
QY 4752 CAACTCGGGAGTGGATCTGTGGCATGGTGCAGTACCAATCGCCCATACAGCTGGTCAG 4693
DB |||||||
QY 361 CGGCTCCAGCACTATGAGTGGATCTTTCGGCAGCCTTGGAAAGTGCAGCAACAGGTGCT 420
DB |||||||
QY 4692 CGGCTCCAGCACTATGAGTGGATCTTTCGGCAGCCTTGGAAAGTGCAGCAACAGGTGCT 4633
DB |||||||
QY 421 CGTTGGCGTACGGAACTGAAATGTACGGCTGCGACGATACGTCGGCGCTGTTGAAATC 480
DB |||||||
QY 4632 CGTTGGCGTACGGAACTGAAATGTACGGCTGCGACGATACGTCGGCGCTGTTGAAATC 4573
DB |||||||
QY 481 ACTCACGAAAAACGGCTGGCAGCTCCACGTTTCAACGGCTTACATGTTTCAAGTGGTAC 540
DB |||||||
QY 4572 ACTCACGAAAAACGGCTGGCAGCTCCACGTTTCAACGGCTTACATGTTTCAAGTGGTAC 4513
DB |||||||
QY 541 GTGAGTGAGAACTCTCGAATCTTTCGGCAGCCTTGGAAAGTGCAGTGGTGCATTCGAA 600
DB |||||||
QY 4512 GTGAGTGAGAACTCTCGAATCTTTCGGCAGCCTTGGAAAGTGCAGTGGTGCATTCGAA 4453
DB |||||||
QY 601 CTGCTATCTCGGATTTGCTGCGGCTTGGTGCAGTACCAATCGCCCATACAGCTGGTCAG 660
DB |||||||
QY 4452 CTGCTATCTCGGATTTGCTGCGGCTTGGTGCAGTACCAATCGCCCATACAGCTGGTCAG 4393
DB |||||||
QY 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGAGCAAAAATTTGCAATCTGGCGTT 720
DB |||||||
QY 4392 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGAGCAAAAATTTGCAATCTGGCGTT 4333
DB |||||||
QY 721 GGTATGGAGGTTGGTGTGGCGACGAAAGTGTGCACATGCAACCGTGCACCCCTGG 780
DB |||||||
QY 4332 GGTATGGAGGTTGGTGTGGCGACGAAAGTGTGCACATGCAACCGTGCACCCCTGG 4273
DB |||||||
QY 781 GAAATCGCTGTTGATGAGTGGCGGATCCACAGCGTTTGAACCTGTGCGGAGAAATTT 840
DB |||||||
QY 4272 GAAATCGCTGTTGATGAGTGGCGGATCCACAGCGTTTGAACCTGTGCGGAGAAATTT 4213
DB |||||||
QY 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGGTGGTCCCGTGGATTCGGTCCCGGAGCT 900
DB |||||||
QY 4212 GAGTTTGGTTCGATGGGACGTCGGGCAATCGGTGGTCCCGTGGATTCGGTCCCGGAGCT 4153
DB |||||||
QY 901 GGTCTTGGGCGAGAACTTAAACAGATGCTCGATCGTGGTCCCGTGGATTCGGTCCCGGAGCT 960
DB |||||||
QY 4152 GGTCTTGGGCGAGAACTTAAACAGATGCTCGATCGTGGTCCCGTGGATTCGGTCCCGGAGCT 4093
DB |||||||
QY 961 ATGGTTGCGATCATTCGGCGCGATCGTGGATGATTCGGATTCGGTCCCGTGGATTCGGTCC 1020
DB |||||||
QY 4092 ATGGTTGCGATCATTCGGCGCGATCGTGGATGATTCGGATTCGGTCCCGTGGATTCGGTCC 4033
DB |||||||
QY 1021 TTCGCGAGATCTTCGGAATCGTGGAGCTGGCGGATCCAGAAATCTTTCGATGATCAC 1080
DB |||||||
QY 4032 TTCGCGAGATCTTCGGAATCGTGGAGCTGGCGGATCCAGAAATCTTTCGATGATCAC 3973
DB |||||||
QY 1081 TTGCATTAATGATTCGCGAGCGATGTCGGGCCCCCGGATTAATTCGGTTCGCAAGTGA 1140
DB |||||||
QY 3972 TTGCATTAATGATTCGCGAGCGATGTCGGGCCCCCGGATTAATTCGGTTCGCAAGTGA 3913
DB |||||||

RESULT 9

US-10-415-562A-7/c
; Sequence 7, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHRL7
US-10-415-562A-7

Query Match 100.0%; Score 1140; DB 17; Length 9652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACAGCTGTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
DB |||||||
QY 5052 ATGACACAGCTGTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 4993
DB |||||||
QY 61 TCCGATTAAGCGCGCATCCGSCACGAACTGCGACCCAAATTTCAACAAATCACCACGTCA 120
DB |||||||
QY 4992 TCCGATTAAGCGCGCATCCGSCACGAACTGCGACCCAAATTTCAACAAATCACCACGTCA 4933
DB |||||||
QY 121 GAAACATTTAACGCTCTGTGGCGGCGGATTTCTGCGTGAACGCTGTGACCAATTTGTCAAC 180
DB |||||||
QY 4932 GAAACATTTAACGCTCTGTGGCGGCGGATTTCTGCGTGAACGCTGTGACCAATTTGTCAAC 4873
DB |||||||
QY 181 GGTCCGAAAGCTTCTGATTCGGAGGCTTCTGCTCGGAAAGGCTGATCTGCCTC 240
DB |||||||
QY 4872 GGTCCGAAAGCTTCTGATTCGGAGGCTTCTGCTCGGAAAGGCTGATCTGCCTC 4813
DB |||||||
QY 241 TGCTGTGCGGAAAGTCTGTCGACATCGTGCAGCAAAATTTCTCAAGTTGTGCTCAT 300
DB |||||||
QY 4812 TGCTGTGCGGAAAGTCTGTCGACATCGTGCAGCAAAATTTCTCAAGTTGTGCTCAT 4753
DB |||||||
QY 301 CAACTCGGGACTGATCTGTTGCGATGCGTGAACGATGACCATGCGCCATACAGCTGGTCAG 360
DB |||||||
QY 4752 CAACTCGGGACTGATCTGTTGCGATGCGTGAACGATGACCATGCGCCATACAGCTGGTCAG 4693
DB |||||||
QY 361 CGGCTCCAGCACTTACGACTGGACTTTTCGGCAGCCTTGGAAAGTGCAGCAACAGGTGCT 420
DB |||||||
QY 4692 CGGCTCCAGCACTTACGACTGGACTTTTCGGCAGCCTTGGAAAGTGCAGCAACAGGTGCT 4633
DB |||||||
QY 421 CGTTGGCGTACGGAACTGAAATGTACGGCTGCGACGATACGTCGGCGCTTGTGAAATC 480
DB |||||||
QY 4632 CGTTGGCGTACGGAACTGAAATGTACGGCTGCGACGATACGTCGGCGCTTGTGAAATC 4573
DB |||||||
QY 481 ACTCACGAAAAACGGCTGGCAGCTCCACGTTTCAACGGCTTACATGTTTCAAGTGGTAC 540
DB |||||||
QY 4572 ACTCACGAAAAACGGCTGGCAGCTCCACGTTTCAACGGCTTACATGTTTCAAGTGGTAC 4513
DB |||||||
QY 541 GTGAGTGAGAACTCTCGAATCTTTCGGCAGCCTTGGAAAGTGCAGTGGTGCATTCGAA 600
DB |||||||
QY 4512 GTGAGTGAGAACTCTCGAATCTTTCGGCAGCCTTGGAAAGTGCAGTGGTGCATTCGAA 4453
DB |||||||
QY 601 CTCGATCTCTGGGATTTGCTGCGCCACTAGCTAAATTCGGGTGGTCTCGATGTAGCAAG 660
DB |||||||
QY 4452 CTCGATCTCTGGGATTTGCTGCGCCACTAGCTAAATTCGGGTGGTCTCGATGTAGCAAG 4393
DB |||||||
QY 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGAGCAAAAATTTGCAATCTGGCGTT 720
DB |||||||
QY 4392 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGAGCAAAAATTTGCAATCTGGCGTT 4333
DB |||||||
QY 721 GGTATGGAGGTTGGTGTGGCGACGAAAGTGTGCACATGCAACCGTGCACCCCTGG 780
DB |||||||
QY 4332 GGTATGGAGGTTGGTGTGGCGACGAAAGTGTGCACATGCAACCGTGCACCCCTGG 4273
DB |||||||
QY 781 GAAATCGCTGTTGATGAGTGGCGGATCCACAGCGTTTGAACCTGTGCGGAGAAATTT 840
DB |||||||
QY 4272 GAAATCGCTGTTGATGAGTGGCGGATCCACAGCGTTTGAACCTGTGCGGAGAAATTT 4213
DB |||||||
QY 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGGTGGTCCCGTGGATTCGGTCCCGGAGCT 900
DB |||||||
QY 4212 GAGTTTGGTTCGATGGGACGTCGGGCAATCGGTGGTCCCGTGGATTCGGTCCCGGAGCT 4153
DB |||||||
QY 901 GGTCTTGGGCGAGAACTTAAACAGATGCTCGATCGTGGTCCCGTGGATTCGGTCCCGGAGCT 960
DB |||||||
QY 4152 GGTCTTGGGCGAGAACTTAAACAGATGCTCGATCGTGGTCCCGTGGATTCGGTCCCGGAGCT 4093
DB |||||||
QY 961 ATGGTTGCGATCATTCGGCGCGATCGTGGATGATTCGGATTCGGTCCCGTGGATTCGGTCC 1020
DB |||||||
QY 4092 ATGGTTGCGATCATTCGGCGCGATCGTGGATGATTCGGATTCGGTCCCGTGGATTCGGTCC 4033
DB |||||||
QY 1021 TTCGCGAGATCTTCGGAATCGTGGAGCTGGCGGATCCAGAAATCTTTCGATGATCAC 1080
DB |||||||
QY 4032 TTCGCGAGATCTTCGGAATCGTGGAGCTGGCGGATCCAGAAATCTTTCGATGATCAC 3973
DB |||||||
QY 1081 TTGCATTAATGATTCGCGAGCGATGTCGGGCCCCCGGATTAATTCGGTTCGCAAGTGA 1140
DB |||||||
QY 3972 TTGCATTAATGATTCGCGAGCGATGTCGGGCCCCCGGATTAATTCGGTTCGCAAGTGA 3913
DB |||||||

Db 4152 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGACGAGAAATCTGCCCGGTC 4093
QY 961 ATGGTTGGGATCATTTCCGGCGGATGCTGGATGATGATTCGGAATTTGGGCTTACGTC 1020
Db 4092 ATGGTTGGGATCATTTCCGGCGGATGCTGGATGATGATTCGGAATTTGGGCTTACGTC 4033
QY 1021 TTCGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTTGGGAAATCTTCGTGATCAC 1080
Db 4032 TTCGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTTGGGAAATCTTCGTGATCAC 3973
QY 1081 TTGCATTATCGATTGCCGACGGATGTCGGCCCCCGGATTAATTCGGTTCCGCAAGTGA 1140
Db 3972 TTGCATTATCGATTGCCGACGGATGTCGGCCCCCGGATTAATTCGGTTCCGCAAGTGA 3913

RESULT 10

US-10-007-527A-6/c

; Sequence 6, Application US/10007527A

; Publication No. US20030044807A1

; GENERAL INFORMATION:

; APPLICANT: Tomb, Jean-Francois

; APPLICANT: Bramucci, Michael G.

; APPLICANT: Cheng, Qiong

; APPLICANT: Kostichka, Kristy N.

; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

; FILE REFERENCE: CL1709 US NA

; CURRENT APPLICATION NUMBER: US/10/007,527A

; CURRENT FILING DATE: 2001-12-05

; PRIOR FILING DATE: 2000-12-05

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6

; LENGTH: 11241

; TYPE: DNA

; ORGANISM: Plasmid prHBR17

US-10-007-527A-6

Query Match

Best Local Similarity 100.0%; Score 1140; DB 14; Length 11241;

Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACGCTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
Db 6641 ATGACACGCTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 6582
QY 61 TCCGATAAGCGCGGCATCCGGCAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 120
Db 6581 TCCGATAAGCGCGGCATCCGGCAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 6522
QY 121 GAACATTTAAGCTGTGGCGCGGATTTCTGGCGTGAACGTTGACCAATTTGTCAC 180
Db 6521 GAACATTTAAGCTGTGGCGCGGATTTCTGGCGTGAACGTTGACCAATTTGTCAC 6462
QY 181 GGTCCGAAAGTTCCTGGAATTCGGAGCTTCTGCTTCCGGAAGGCTGGATTCGCC 240
Db 6461 GGTCCGAAAGTTCCTGGAATTCGGAGCTTCTGCTTCCGGAAGGCTGGATTCGCC 6402
QY 241 TGCTGTGGGAAAGTTCGGTGCACATCGTGACAGCAAAATTTCTCAAGTTGTTGCTCAT 300
Db 6401 TGCTGTGGGAAAGTTCGGTGCACATCGTGACAGCAAAATTTCTCAAGTTGTTGCTCAT 6342
QY 301 CAACCTCGGAGTTCGATCTGTTGCGATGTTGACGATGACCATGCGCATACAGCTGTCAG 360
Db 6341 CAACCTCGGAGTTCGATCTGTTGCGATGTTGACGATGACCATGCGCATACAGCTGTCAG 6282
QY 361 CGGCTCCACGCTATGACATGGAATTTTCGGCAGCTTCCGCAAGCTGCGCAACCGTTCGT 420
Db 6281 CGGCTCCACGCTATGACATGGAATTTTCGGCAGCTTCCGCAAGCTGCGCAACCGTTCGT 6222
QY 421 CGTTGGGCTACGGAACGTAATGATGACGCTTCGACGATACGTCGCGCTGTGGAATC 480

Db 6221 CGTTGGGCTACGGAACGTAATGATACGCTTCGACGGATACGTCGCGCTGTGAAATC 6162
QY 481 ACTACGGAATAAAACGGCTGGCACTGCACTGTTACGGCTACTCATGTTAGTGGTGAC 540
Db 6161 ACTACGGAATAAAACGGCTGGCACTGCACTGTTACGGCTACTCATGTTAGTGGTGAC 6102
QY 541 GTGAGTGAGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGGACTTCCAA 600
Db 6101 GTGAGTGAGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGGACTTCCAA 6042
QY 601 CTCGTATCTCTGGGATTTGCTGCGCACCTACGTAATTCGGGTGGTCTCGATACGAAAG 660
Db 6041 CTCGTATCTCTGGGATTTGCTGCGCACCTACGTAATTCGGGTGGTCTCGATACGAAAG 5982
QY 661 ATCGCGGTGAAAGCTGATCAAGTTCTCGTGGTATCTGAACGAAATTCGATCTGGCGTT 720
Db 5981 ATCGCGGTGAAAGCTGATCAAGTTCTCGTGGTATCTGAACGAAATTCGATCTGGCGTT 5922
QY 721 GGTATGGAGGTTGGTAGTGGGACGAAAGTGTGACATGGCAACCGTGCACCCCTGG 780
Db 5921 GGTATGGAGGTTGGTAGTGGGACGAAAGTGTGACATGGCAACCGTGCACCCCTGG 5862
QY 781 GAAATCGCTGTTGATGCAAGTGGCGGGATCCACAAGCGTTTGAATCTGTGGCGAATTT 840
Db 5861 GAAATCGCTGTTGATGCAAGTGGCGGGATCCACAAGCGTTTGAATCTGTGGCGAATTT 5802
QY 841 GAGTTGGTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGGATTCGCTGCCCGAGCT 900
Db 5801 GAGTTGGTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGGATTCGCTGCCCGAGCT 5742
QY 901 GGTCTTGGGCGAGACTTACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAGACTTACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 5682
QY 961 ATGTTGCGATCATTTCCGCGCGATCGTGATGATGATTCGGAATCTGTGGCTTACGTC 1020
Db 5681 ATGTTGCGATCATTTCCGCGCGATCGTGATGATGATTCGGAATCTGTGGCTTACGTC 5622
QY 1021 TTCGCGAGATCTTCGGAATCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 1080
Db 5621 TTCGCGAGATCTTCGGAATCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562
QY 1081 TTGATTTGATTCGCTCCGCGAGCGATGTCGGCGCCCGATTAATTCGTTTCGCAAGTGA 1140
Db 5561 TTGATTTGATTCGCTCCGCGAGCGATGTCGGCGCCCGATTAATTCGTTTCGCAAGTGA 5502

RESULT 11

US-10-007-452-6/c

; Sequence 6, Application US/10007452

; Publication No. US20030093701A1

; GENERAL INFORMATION:

; APPLICANT: Tomb, Jean-Francois

; APPLICANT: Bramucci, Michael G.

; APPLICANT: Cheng, Qiong

; APPLICANT: Kostichka, Kristy N.

; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

; FILE REFERENCE: CL1709 US NA

; CURRENT APPLICATION NUMBER: US/10/007,452

; CURRENT FILING DATE: 2001-11-08

; PRIOR APPLICATION NUMBER: 60/254,868

; PRIOR FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6

; LENGTH: 11241

; TYPE: DNA

; ORGANISM: Plasmid prHBR17

US-10-007-452-6

Query Match

Best Local Similarity 100.0%; Score 1140; DB 14; Length 11241;

Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGACACGCTAGTGTCTGAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGTGTCG 60
Db |||||||
6641 ATGACACGCTAGTGTCTGAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGTGTCG 6582

Qy 61 TCCGATAAGCGCGCATCCCGCAGCAAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 120
Db |||||||
6581 TCCGATAAGCGCGCATCCCGCAGCAAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 6522

Qy 121 GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCAATTTGTCAC 180
Db |||||||
6521 GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCAATTTGTCAC 6462

Qy 181 GGTCCGAAAGTTCGTGATTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 240
Db |||||||
6461 GGTCCGAAAGTTCGTGATTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 6402

Qy 241 TGCTGTCCGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db |||||||
6401 TGCTGTCCGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 301 CAATCCGGGACTGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 360
Db |||||||
6341 CAATCCGGGACTGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 6282

Qy 361 CGGCTCCAGACCTATGGAAGTTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 420
Db |||||||
6281 CGGCTCCAGACCTATGGAAGTTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 6222

Qy 421 CGTTGGCGTACCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 480
Db |||||||
6401 TGCTGTCCGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 481 ACTCAGCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 540
Db |||||||
6161 ACTCAGCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 6102

Qy 541 GTGAGTGAAGATCTCTGGAATTCGCGGCTGCTGATGAGTGAAGTTCGATCGGTCGAAATC 600
Db |||||||
6101 GTGAGTGAAGATCTCTGGAATTCGCGGCTGCTGATGAGTGAAGTTCGATCGGTCGAAATC 6042

Qy 601 CTGCTATCTCTGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 660
Db |||||||
5981 ATCGCGGCTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTTGCAATCTCGGCTT 5922

Qy 721 GGTATGAGGTTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 780
Db |||||||
5921 GGTATGAGGTTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5862

Qy 781 GAAATCGCTGTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 840
Db |||||||
5861 GAAATCGCTGTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5802

Qy 841 GAGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 900
Db |||||||
5801 GAGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5742

Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTCGAGGAGGAGATCTGCCCCGGTC 960
Db |||||||
5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTCGAGGAGGAGATCTGCCCCGGTC 5682

Qy 961 ATGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 1020
Db |||||||
5681 ATGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5622

Qy 1021 TTCGGCGAGATCTCTCGGACTCGTGCAGACCTGCGCGGAGTTCGAGGAAATCTTTCGTCGATCAC 1080
Db |||||||
5621 TTCGGCGAGATCTCTCGGACTCGTGCAGACCTGCGCGGAGTTCGAGGAAATCTTTCGTCGATCAC 5562
```

```
Qy 1081 TTGCATTATCGATTGCCCGCAGCGGATGTGGGCCCCCGGATAATATATCGTTCGCAAGTGA 1140
Db |||||||
5561 TTGCATTATCGATTGCCCGCAGCGGATGTGGGCCCCCGGATAATATCGTTCGCAAGTGA 5502

RESULT 12
US-10-415-562A-6/c
; Sequence 6, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: Cl1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid_pRHBR17
US-10-415-562A-6

Query Match 100.0%; Score 1140; DB 17; Length 11241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACACGCTAGTGTCTGAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGTGTCG 60
Db |||||||
6641 ATGACACGCTAGTGTCTGAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGTGTCG 6582

Qy 61 TCCGATAAGCGCGCATCCCGCAGCAAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 120
Db |||||||
6581 TCCGATAAGCGCGCATCCCGCAGCAAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 6522

Qy 121 GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCAATTTGTCAC 180
Db |||||||
6521 GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCAATTTGTCAC 6462

Qy 181 GGTCCGAAAGTTCGTGATTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 240
Db |||||||
6461 GGTCCGAAAGTTCGTGATTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 6402

Qy 241 TGCTGTCCGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db |||||||
6401 TGCTGTCCGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 301 CAATCCGGGACTGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 360
Db |||||||
6341 CAATCCGGGACTGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 6282

Qy 361 CGGCTCCAGACCTATGGAAGTTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 420
Db |||||||
6281 CGGCTCCAGACCTATGGAAGTTCGGAGGCTTCGTTCCTGCGGAAAGGCTGATCGCCCC 6222

Qy 421 GGTTCGCGTACCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 480
Db |||||||
6221 GGTTCGCGTACCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 6162

Qy 481 ACTCAGCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 540
Db |||||||
6161 ACTCAGCGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 6102

Qy 541 GTGAGTGAAGATCTCTGGAATTCGCGGCTGCTGATGAGTGAAGTTCGATCGGTCGAAATC 600
Db |||||||
6101 GTGAGTGAAGATCTCTGGAATTCGCGGCTGCTGATGAGTGAAGTTCGATCGGTCGAAATC 6042

Qy 601 CTGCTATCTCTGGGAAAGTTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 660
Db |||||||
5981 ATCGCGGCTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTTGCAATCTCGGCTT 5922

Qy 721 GGTATGAGGTTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 780
Db |||||||
5921 GGTATGAGGTTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5862

Qy 781 GAAATCGCTGTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 840
Db |||||||
5861 GAAATCGCTGTGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5802

Qy 841 GAGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 900
Db |||||||
5801 GAGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5742

Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTCGAGGAGGAGATCTGCCCCGGTC 960
Db |||||||
5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTCGAGGAGGAGATCTGCCCCGGTC 5682

Qy 961 ATGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 1020
Db |||||||
5681 ATGTTTGGTTCGATGAGGCGGAGCGGAAAGTTCGATGAGTGAAGTTCGATCGGTCGAAATC 5622

Qy 1021 TTCGGCGAGATCTCTCGGACTCGTGCAGACCTGCGCGGAGTTCGAGGAAATCTTTCGTCGATCAC 1080
Db |||||||
5621 TTCGGCGAGATCTCTCGGACTCGTGCAGACCTGCGCGGAGTTCGAGGAAATCTTTCGTCGATCAC 5562
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 02:18:21 ; Search time 85 Seconds
(without alignments)
1599.511 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005
Sequence: 1 MTSVSAHLSGKDRPPVLVS.....HLHYRLPAADVRPPIISVRK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2005	100.0	379	5	ABB84278
2	609.5	30.4	459	5	ABB84280
3	399.5	19.9	494	2	AAW73071
4	326	16.3	427	3	AAW97536
5	326	16.3	427	5	ABB06341
6	326	16.3	427	5	AAE13530
7	322	16.1	451	5	ABB84283
8	318	15.9	427	3	AAW97537
9	318	15.9	427	5	ABB06342
10	318	15.9	427	5	AAE13531
11	318	15.9	427	7	ADB66179
12	316	15.8	456	5	ABB84281
13	282	14.1	466	5	AAU11039
14	266.5	13.3	528	5	ABB84282
15	134	6.7	314	1	AAW93143
16	104	5.2	409	6	AAE37298
17	103	5.1	647	6	ABU56572
18	101.5	5.1	882	7	ABO68907
19	101	5.0	1395	7	ABO81599
20	100.5	5.0	577	2	AAW06554
21	100	5.0	1568	4	ABG03109
22	99.5	5.0	676	7	ABO74435
23	98.5	4.9	532	6	ABU15733
24	98.5	4.9	621	7	ABO70485
25	98.5	4.9	1409	7	ABO73536

ALIGNMENTS

RESULT 1
ABB84278

ID ABB84278 standard; protein; 379 AA.

XX ABB84278;

XX AC ABB84278;

DT 13-JAN-2003 (first entry)

DE Rhodococcus AN12 replication protein Rep.

XX Rhodococcus AN12 replication protein Rep.

KW Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector; Rep protein.

XX Rhodococcus erythropolis.

XX WO200255709-A2.

XX 18-JUL-2002.

XX 12-DEC-2001; 2001WO-US047868.

XX 12-DEC-2000; 2000US-0254868P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

XX WPI; 2002-557827/59.

XX N-PSDB; ABQ76122.

XX New nucleic acid molecule encoding replication protein/plasmid stability protein, useful in cloning and expression vectors, particularly shuttle vectors for expression of heterologous genes in Rhodococcus species.

XX Claim 4; Page 64-65; 96pp; English.

XX This invention describes a novel nucleic acid encoding a replication protein or a plasmid stability protein. The product of the invention is useful for expression of nucleic acid such as genes encoding enzymes involved in the production of isoprenoid molecules, polyhydroxyalkanoic acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol dehydrogenase, terpene synthases, and cholesterol oxidase in an Actinomycetales bacteria. The replication protein or plasmid stability protein are useful in cloning and expression vectors and particularly in

CC	shuttle vectors for the expression of homologous and heterologous genes
CC	in <i>Rhodococcus</i> sp. This sequence represents the <i>Rhodococcus</i> AN12 Rep
CC	protein described in the disclosure of the invention
XX	
SQ	Sequence 379 AA;
	Query Match 100.0%; Score 2005; DB 5; Length 379;
	Best Local Similarity 100.0%; Pred. No. 6.6e-193;
	Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MTVSASAEHLGSKORPPVLVSSDKGI RHLELPKLCQITTSFTFNACGRPIGVNGVTIVN 60
DB	1 MTVSASAEHLGSKORPPVLVSSDKGI RHLELPKLCQITTSFTFNACGRPIGVNGVTIVN 60
QY	61 GPKGSGFGGLRSCKGKICPCACAGVGAAHRADEISQVVAHQLTGTSVAMVTWTRHTTAGQ 120
DB	61 GPKGSGFGGLRSCKGKICPCACAGVGAAHRADEISQVVAHQLTGTSVAMVTWTRHTTAGQ 120
QY	121 RLHDLTWLGTLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGMHVHVALLMFSGD 180
DB	121 RLHDLTWLGTLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGMHVHVALLMFSGD 180
QY	181 VSENILESFSDAMDRTWTSKLVLGFAPLPLNSGGLDVRKTCGGBADQVLAAYLTAKIAGSV 240
DB	181 VSENILESFSDAMDRTWTSKLVLGFAPLPLNSGGLDVRKTCGGADQVLAAYLTAKIAGSV 240
QY	241 GMEVSGDGKSGRGHNAPWIEAVDVAGGDPOALELMREFEFGSMGRGAIAWSRGLRARA 300
DB	241 GMEVSGDGKSGRGHNAPWIEAVDVAGGDPOALELMREFEFGSMGRGAIAWSRGLRARA 300
QY	301 GLGAELTDAQIVEOEESAPVMVAIIIPARSWMMIRTCAFYVFEGEILGLVEAGATWENLRDH 360
DB	301 GLGAELTDAQIVEOEESAPVMVAIIIPARSWMMIRTCAFYVFEGEILGLVEAGATWENLRDH 360
QY	361 LHYRLPAADVRRPPTIISVRK 379
DB	361 LHYRLPAADVRRPPTIISVRK 379

RESULT 2	
ABB84280	
ID	ABB84280 standard; protein; 459 AA.
XX	
AC	ABB84280;
XX	
DT	13-JAN-2003 (first entry)
XX	
DE	A. pyogenes pAP1 replication protein.
XX	
KW	Plasmid stability protein; replication protein; ethylene forming enzyme;
KW	carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW	polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW	alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW	shuttle vector.
XX	
OS	Arcanobacterium pyogenes.
XX	
PN	WO200255709-A2.
XX	
PD	18-JUL-2002.
XX	
PF	12-DEC-2001; 2001WO-US047868.
XX	
PR	12-DEC-2000; 2000US-0254868P.
XX	
PA	(DUPO) DU PONT DE NEMOURS & CO E I.
XX	
PI	Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX	
DR	WPI; 2002-557827/59.
XX	
PT	New nucleic acid molecule encoding replication protein/plasmid stability
PT	protein, useful in cloning and expression vectors, particularly shuttle
PT	

[illegible]

Db 140 RARGVHTYSDYEVTSWANGHLHRNMLFLDRPLSDDELKAFESMFSRWAGVYKA 199
 Qy 205 GFAAPLRNSG-GLD-VRKIGEADQVLAAYLTKTASGVGMVSGDGKSGRHNAPWEI 262
 Db 200 GMDAPLREHGKLVDOVSTWGGDAK-MATYLAK---GMSQELTGSATKTSKSGVTPFQM 255
 Qy 263 AVDAVGDDPQALE-----LWREFFEGSGRAIAWSGRIGARAGIGABLTDAQI--- 311
 Db 256 -LDMLAQSDAGEDMDAVLVARWEYEVGSKNLRN-SWSRG--AKRALGIDYIDAVRRE 311
 Qy 312 VEQE-----ESAPVMVAITPARSMMIRT 335
 Db 312 MEELYKLAGLEAPERVESTVAVALYKPDWKLIQS 348

RESULT 7
 ABB84283
 ID ABB84283 standard; protein; 451 AA.
 AC ABB84283;
 XX
 DT 06-AUG-2003 (revised)
 DT 13-JAN-2003 (first entry)
 XX
 DE S. nigrificans pSN22 replication protein.
 XX
 KW Plasmid stability protein; replication protein; ethylene forming enzyme;
 KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
 KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
 KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
 KW shuttle vector.
 XX
 OS Streptomyces nigrescens.
 XX
 FN WO200255709-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 12-DEC-2001; 2001WO-US047868.
 XX
 PR 12-DEC-2000; 2000US-0254868P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 FI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
 XX
 DR WPI; 2002-557827/59.
 XX
 PT New nucleic acid molecule encoding replication protein/plasmid stability
 PT protein, useful in cloning and expression vectors, particularly shuttle
 PT vectors for expression of heterologous genes in Rhodococcus species.
 XX
 PS Example 5; Fig 4A; 96pp; English.
 CC
 CC This invention describes a novel nucleic acid encoding a replication
 CC protein or a plasmid stability protein. The product of the invention is
 CC useful for expression of nucleic acid such as genes encoding enzymes
 CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
 CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
 CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
 CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
 CC Actinomycetales bacteria. The replication protein or plasmid stability
 CC protein are useful in cloning and expression vectors and particularly in
 CC shuttle vectors for the expression of homologous and heterologous genes
 CC in Rhodococcus sp. This sequence represents the S. nigrificans pSN22
 CC replication protein described in the disclosure of the invention.
 CC (Updated on 06-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 451 AA;

Query Match 16.1%; Score 322; DB 5; Length 451;
 Best Local Similarity 30.8%; Pred. No. 3.2e-23;

Matches 101; Conservative 41; Mismatches 110; Indels 76; Gaps 15;
 Qy 49 PISGV-----NGVTIYNGPKSGFGGLRSCGKWI CPCACGKGAHRAHISQVVAHQ 102
 Db 3 PASGVI VAQTAAGTSVVL-----GLMRCGRITWLCFVCAATIRHKRAEBITAAVVEWI 54
 Qy 103 GTGSVA-MVTMTMHTAGQRLHDLWTGLSAAWKAATNGRR-----WRTE--- 145
 Db 55 KRGGTAYLVITFTARHGHTRDLADLMDALQGRKTDADAPRRPGAYQRLITGGTWAGRRAKD 114
 Qy 146 -----REMYGCDGYVRAVEITHGK-NGMHVHVHALMFSGDV-----SENILBSF 189
 Db 115 GHRAADREGIRDRIGYVGMIRATEVTVGQINGWPHIHAIVLVGRTGERSAQIVGTF 174
 Qy 190 --SDAMFDRWTSKLVSLGFAPIR-----NSGGLDVRKIGIHEAD-QVLAAYLTAKIA 237
 Db 175 EPSEALDEWQGWRAV-WTAALKRVNPQFTPDDRHGVDFKRLETERDANDLABYIAKTQ 233
 Qy 238 SG--VGMEVSGDGKSGRHNAPWEI--AVDAVG--DPQA-----LELWREFE 281
 Db 234 DGKAPALELADLKTANGNVAFELLGRIGLDTGGWTEDDAAGVGSLEWNLARWHEYE 293
 Qy 282 FGSNGRAIAWSRIGRARAGLGABLTDA 309
 Db 294 RATKGRRAIEWTRYLRQMLGLDGGDTEA 321

RESULT 8
 AAY97537
 ID AAY97537 standard; protein; 427 AA.
 XX
 AC AAY97537;
 XX
 DT 12-SEP-2003 (revised)
 DT 15-JAN-2001 (first entry)
 XX
 DE B. lactofermentum pSFK6 protein sequence.
 XX
 KW Temperature sensitive plasmid; TSRCR; protein production;
 KW temperature sensitive replication control region; pSFK6.
 XX
 OS Corynebacterium glutamicum.
 XX
 FN EPI038966-A1.
 XX
 PD 27-SEP-2000.
 XX
 PF 16-MAR-2000; 2000EP-00105326.
 XX
 PR 16-MAR-1999; 99JP-00069896.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 FI Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;
 XX
 DR WPI; 2000-573832/54.
 DR N-PSDB; AAA90951.
 XX
 PT Plasmids containing a temperature sensitive replication control regions
 PT useful for breeding microorganisms for the production of amino acids by
 PT fermentation.
 XX
 PS Disclosure; Page 19-22; 29pp; English.
 XX
 CC This sequence is the Brevibacterium lactofermentum pSFK6 protein. The
 CC invention relates to a plasmid containing a temperature sensitive
 CC replication control region (TSRCR) and a marker gene (MG). The TSRCR is
 CC derived from plasmid pAM330 harboured by Brevibacterium lactofermentum
 CC (ATCC 13869) and allows the plasmid to replicate autonomously at a low
 CC temperature but does not allow the plasmid to replicate autonomously at
 CC an elevated temperature in coryneform bacteria within a temperature range
 CC in which the bacteria can grow. The plasmid can be used for modifying a
 CC chromosomal gene in a coryneform bacterium, which may be used for the

XX Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector.
XX Streptomyces lividans.
OS WO200255709-A2.
FN PD 18-JUL-2002.
XX PF 12-DEC-2001; 2001WO-US047868.
XX PR 12-DEC-2000; 2000US-0254868P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PA Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX WPI; 2002-557827/59.
XX New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.
XX Example 5; Fig 4A; 96pp; English.
XX This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence represents a Rhodococcus AN12 Rep
CC protein described in the disclosure of the invention
XX Sequence 456 AA;
Query Match 15.8%; Score 316; DB 5; Length 456;
Best Local Similarity 30.5%; Pred. No. 1.3e-22;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;
QY 49 PTSGV-----NGVTIVNGPKSGFGGLRSCKGKWCPCCKAGVGAHRADEISQVVAHQ 102
Db 3 PASGVIVAQTAAAGTSVVL-----GLMRCGRILWCPVCAATIRKRAEITAAVVEWI 54
QY 103 GTGSA-VYTMTRHTAGORLHDLWTGLSAAWKAATNGRR-----WRTE----- 145
Db 55 KRGGTAYLVTFARHGHTDRDLADLMDALQGRKTPDSPRPGAYQRLTGGTWAGRRAKD 114
QY 146 -----REMYCDGVVRAVEITHGK-NGWVHVHALLMFSGDV-----SENILESF 189
Db 115 GHRAADRGRIDRIGVGMIRATEVTVGQINGWPHIHAIVLGVGRTGERSAKQIVATF 174
QY 190 --SDAMFRWTSKLVSLGFAAPLR-----NSGGLDVRIKIGGEAD-QVLAAYLTAKIA 237
Db 175 EPTGAALDEWQGHWSV-WTAALRKVNPAFTDDRHGVDFKLETERDANDLAEYIAKIQ 233
QY 238 SG--VGMVEGSDGSGRHNAPWEI---AVDVG-----DPOA-----LELWREPE 281
Db 234 DGKAPALELARADLTATGNNVAPPELLGRIGDLTGGMTEDDAAGVGSLEWNLRSRWEYE 293
QY 282 FSGMRRATAWSRGLRARAGLGAELTDA 309
Db 294 RATRGRRALEWTRYLRQMLGLDGGDTEA 321

RESULT 13
AAU11039
ID AAU11039 standard; protein; 466 AA.
XX AC AAU11039;
XX 29-AUG-2003 (revised)
DT 07-AUG-2003 (revised)
DT 12-MAR-2002 (first entry)
XX DE
XX Replication protein encoded by Ketogulonigenium plasmid pADMX6L2.
KW Ketogulonigenium; 2-keto-L-gulononic acid; L-sorbose; sorbitol;
KW replication protein; plasmid pADMX6L2.
XX OS
XX Ketogulonigenium robustum.
XX PN WO200177348-A2.
XX PD 18-OCT-2001.
XX PF 05-APR-2001; 2001WO-US011097.
XX PR 05-APR-2000; 2000US-0194627P.
XX PA (ARCH) ARCHER-DANIELS MIDLAND CO.
PA (DELI) D'ELIA J.
PA (STOD) STODDARD S F.
XX D'elia J, Stoddard SF;
XX WPI; 2002-041295/05.
DR N-PSDB; AAS17120.
XX New bacterium of Ketogulonigenium genus, useful for producing 2-keto-L-
PT gulonic acid from sorbose or sorbitol, comprises transgene containing DNA
PT sequence from endogenous Ketogulonigenium plasmid.
XX Disclosure; Fig 9; 116pp; English.
XX The present invention relates to a new bacterium of genus
CC Ketogulonigenium. Ketogulonigenium may further comprise a transgene,
CC comprising a DNA sequence from an endogenous Ketogulonigenium plasmid.
CC Methods for transforming Ketogulonigenium are also described. The
CC invention is useful for producing 2-keto-L-gulononic acid (2-KLG) from L-
CC sorbose or sorbitol. The present sequence represents the replication
CC protein encoded by the Ketogulonigenium endogenous plasmid pADMX6L2.
CC (Updated on 07-AUG-2003 to correct OS field.) (Updated on 29-AUG-2003 to
CC standardise OS field)
XX Sequence 466 AA;
Query Match 14.1%; Score 282; DB 5; Length 466;
Best Local Similarity 30.0%; Pred. No. 3.5e-19;
Matches 87; Conservative 40; Mismatches 119; Indels 44; Gaps 15;
QY 60 NGPKSGFGGLRSCKGKWCPCCKAGVGAHRADEISQVVAHQLGTSV-AMVTMTMRHTA 118
Db 109 DGGREAFFSGGLAQCRNVWGCVCASARIAQIRSEMNHLLAWARDNGFVPLVLTIAQHK 168
QY 119 GORLHDLWTGLSAAWKAATNGRRWRTEREMYGCD-----GYVRAVEITHG-KNGWVHVH 173
Db 169 GDSLFDDLQNMKKAQRLRQRREW-----DLPFVGSVTSTSTIHSYANGWHPHFHE 220
QY 174 L-LMFSGDVSE--NILESFDAMFDRWTSKLVSLGFAAPLRNSGGLDVRIKIGGEADQVLA 230
Db 221 IYLLRAGDESEALHLMQRLGDA---WRACLKGYGMWG---NDAAFDVRGAANAGD----- 269
QY 231 AYLTKIASGVGMVEGSGDKSGRHNAPWEIYAVDVGDPQALBLWREFFEGSMG--RR 288
Db 270 -YVAK--WGAABELTSSSKGRKGRTPRQLL--QAGDD-----GLWLEYFNATSGKRR 320
QY 289 AIAWSRGLRARAGLGAELTD-----AQIVEQESAFVMVAIIPARSMIMIR 334

SQ Sequence 314 AA;

Query Match	6.7%;	Score 134;	DB 1;	Length 314;
Best local Similarity	22.4%;	Pred. No. 0.00017;		
Matches 56;	Conservative 50;	Mismatches 106;	Indels 38;	Gaps 10;
Qy	75	KGMICPCAGKGVGAHRADEISQVV--AH-QLGTGSGVAMVTMRHTAGORLHDLWTGLSA	131	
Db	69	KSRLCPLCNWRRSMGQSNQLMQVLDEAHKQRTKTRFLFLTLTAENASGNLQVEKRWGR	128	
Qy	132	AWKAATNGRWRTREMYGCDGYVRAVETIHKNG-KWHVHVHALLMESGDVSENILESFS	190	
Db	129	A-----ISKLFQYKKPKAKNLLGYVRSFTEITKNGTTHQEHMVLLFVKPYFKDSANYIN	183	
Qy	191	DAMFDRTSKLIVSLGFAPLRNSGGLDVRIKGESADQVLAAYLTKIASGVMEVSGDGK	250	
Db	184	DAEWSKULXWRAMKLDY-XPIVWVAVRSNKAAGNSLIASAQET-----	226	
Qy	251	SGRHGNFAPWEI-AVDVAGGDPOA-LELWREFEFGSMGRRALAWSRGL--RARAGLGAE	306	
Db	227	-----AKYQVSKDILITNDQERDLQWVEDLEQLAGSRQISYG-GLFKEIRKQLQLED	278	
Qy	307	TDQIVQEQE	316	
Db	279	VDAMLINVD	288	

Search completed: October 23, 2004, 03:55:45
Job time : 88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 03:43:51 ; Search time 26 Seconds
(without alignments)
966.712 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGSKDPPVLVS.....HLHVELPAADVRPIISVRK 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	326	16.3	427	3	US-09-521-668B-18
2	318	15.9	427	3	US-09-521-668B-20
3	101.5	5.1	882	4	US-09-252-991A-17653
4	101	5.0	1395	4	US-09-252-991A-30345
5	99.5	5.0	676	4	US-09-252-991A-23181
6	98.5	4.9	621	4	US-09-252-991A-19231
7	98.5	4.9	1409	4	US-09-252-991A-22282
8	97.5	4.9	1245	4	US-09-252-991A-30935
9	97	4.8	713	3	US-09-335-409-11
10	97	4.8	713	3	US-09-568-102-11
11	97	4.8	713	3	US-09-567-969-11
12	97	4.8	713	3	US-09-568-480-11
13	97	4.8	713	3	US-09-568-486-11
14	97	4.8	713	3	US-09-568-472-11
15	97	4.8	713	3	US-09-567-899-11
16	97	4.8	1403	4	US-09-252-991A-25500
17	96.5	4.8	309	4	US-09-489-039A-10229
18	95.5	4.8	626	4	US-09-252-991A-30863
19	95.5	4.8	2890	3	US-09-413-814-67
20	95.5	4.8	3798	3	US-09-335-409-6
21	95.5	4.8	3798	3	US-09-568-102-6
22	95.5	4.8	3798	3	US-09-567-969-6
23	95.5	4.8	3798	3	US-09-568-480-6
24	95.5	4.8	3798	3	US-09-568-486-6
25	95.5	4.8	3798	3	US-09-568-472-6
26	95.5	4.8	3798	3	US-09-567-899-6
27	94.5	4.7	429	4	US-09-252-991A-31382

28	94	4.7	414	4	US-09-252-991A-22296	Sequence 22296, A
29	93.5	4.7	559	4	US-09-489-039A-9319	Sequence 9319, Ap
30	92.5	4.6	579	4	US-09-198-452A-918	Sequence 918, App
31	92.5	4.6	612	4	US-09-252-991A-29283	Sequence 29283, A
32	92	4.6	671	4	US-09-252-991A-31862	Sequence 31862, A
33	91.5	4.6	430	4	US-09-489-039A-10370	Sequence 10370, A
34	91.5	4.6	574	4	US-09-489-039A-13400	Sequence 13400, A
35	91	4.5	158	4	US-09-252-991A-16612	Sequence 16612, A
36	91	4.5	561	4	US-09-252-991A-16726	Sequence 16726, A
37	91	4.5	1174	4	US-09-252-991A-29279	Sequence 29279, A
38	90.5	4.5	611	4	US-09-252-991A-25119	Sequence 25119, A
39	89.5	4.5	415	4	US-09-252-991A-25764	Sequence 25764, A
40	89.5	4.5	601	4	US-09-252-991A-30589	Sequence 30589, A
41	89	4.4	503	4	US-09-252-991A-23522	Sequence 23522, A
42	89	4.4	881	4	US-09-252-991A-31702	Sequence 31702, A
43	89	4.4	3562	4	US-09-679-279-14	Sequence 14, Appl
44	89	4.4	3739	3	US-09-320-878-2	Sequence 2, Appl
45	89	4.4	3739	3	US-09-105-537-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-521-668B-18
; Sequence 18, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, ETICHIRO
; APPLICANT: MATSUI, KAUHIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521.668B
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: JP 11-69896
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 427
; TYPE: EXT
; ORGANISM: Brevibacterium lactofermentum
US-09-521-668B-18

Query Match		16.3%	Score 326;	DB 3;	Length 427;
Best Local Similarity		29.1%	Pred. No. 5.6e-25;		
Matches		98;	Conservative 58;	Mismatches 131;	Indels 50; Gaps 15;
Qy	34	LOQITSETFNACGR-----PISGVNGVTIVNGKSGFGGLRSCGKGWICPCCKAGV	86		
Db	27	MYKITNSKALAGCHRRRDEAVVSS-----NG--ASQFEGQLNSHRSWGSPLAELV	79		
Qy	87	GAHRADISQVAHQLTGTVAMVTM-TMRHTAQRLHDLMTGLSAKAAATNGRRWTE	145		
Db	80	MGERRIELATKKNHLAGGALAMFVGTVRHRNSQSPAQVEAGIKTAYSSMVKTSQWKE	139		
Qy	146	REMYGCDGYVRAVEITHG-KNGMHHVHALLMFGSDYSENILSFSDAMFDRWTSKLVS	204		
Db	140	RARYGVHTYSDYEVTSWANGHLHRNMLLFDRPLSDDELKAFEDSMFSRWAGVVKA	199		
Qy	205	GFAAPLRNSG-GLD-VRKIGGEADQVLAAYITKIASGVMEVSGDGKSGRHGNRAPEWI	262		
Db	200	GMDAPLREHGKLDQVSTWGDAAK-NATYIAK---GMSQELTGSATKTSKGYTTFQM	255		
Qy	263	AVDAVGGDPALE-----LWREFEFGSMGRRAIWSGLRAGIAGELTDAQI---	311		
Db	256	-LQMLADQSDAGEDMDAVLVARWREYEVGSKNLS-SWSRG--AKRALGIDYIDAVRRE	311		
Qy	312	VEQE-----ESAPVMVAITPARSWMIRT	335		

Db 312 MEEELYKLAGLEAPERVESTRAVALVKPDDMKLIQS 348

RESULT 2

US-09-521-668B-20
; Sequence 20, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAUKIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521,668B
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: JP 11-69896
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-521-668B-20

Query Match 15.9%; Score 318; DB 3; Length 427;
Best Local Similarity 28.8%; Pred. No. 3.7e-24;
Matches 97; Conservative 58; Mismatches 132; Indels 50; Gaps 15;

QY 34 LQQTTSFTFACGR-----PISGVNGVTIVNGPKSGRGLRSCGKGCWICPCCAKV 86
DB 27 MYKITNSKALAGCHRRRDEAVASWS-----NG--ASQFEGJONSHRWSGSSIAELEV 79
QY 87 GAHRADEISQVVAHQLTGTSVAMVTM-TMRHTAGQRLHDLMTGLSAAKAAATNGRRWRTE 145
DB 80 MGERRIELAIATKNHLAAGGALMEFVGIVRNRSSQFPAQVAGIKTAYSSVKTSQWKKE 139
QY 146 REMTGCDCGVRAVELTHG-KNGWHVHALLMFGSDVSENILESPSDAMFDRWTSKLVS 204
DB 140 RARYGVEHTYSDYEVTDWANGWHLHRNMLFLDRPLSDDELKAPEDSMFSRWSAGVYKA 199
QY 205 GFAAPLRNSG-GLD-VRKIGEADQVLAAYLTKTASGVGMVSGDGKSGRHRNAPWEI 262
DB 200 GMDAPLRHGVKLDQVSTWGGDAK-MATYIAK-----GMSQELTGSATXATASKGSTYTFQM 255
QY 263 AVDAVGDDPOALE-----LWREFEFGSMRRRAIWSRGLRAGLGAELTDAQI--- 311
DB 256 -LDMLAQSDAGEDMDAVLVARWREYEVGSKNLR-SWSRG-AKRALGIDYIDADVRE 311
QY 312 VEQE-----ESAPVMAIIPARSWMMIRT 335
DB 312 MEEELYKLAGLEAPERVESTRAVALVKPDDMKLIQS 348

RESULT 3

US-09-252-991A-17653
; Sequence 17653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17653

; LENGTH: 882

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (730)

; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-17653

Query Match 5.1%; Score 101.5; DB 4; Length 882;
Best Local Similarity 21.7%; Pred. No. 0.26;
Matches 87; Conservative 48; Mismatches 125; Indels 141; Gaps 20;

QY 47 GRPISGVNGVTIVNGPKSGRGLRSCGKGCWICPCA--GKVGAAHRADEISQVV----- 98
DB 487 GRVLGGVIGLALRQCAIGNR-----VAPFVAQAGRKVGLEAAEQLSRPMHRPTE 535
QY 99 -----AHQLGTGS-----VAMVTMTMRHTAGQRLHDL-----W 126
DB 536 IRHAPADGVGGGAGRRAVELGGGGQQLRLLVVQVRIAL-HDQQRAGGVGRGHRGTGL 594
QY 127 TGLSAAWKAATNGRRWRTEREMYCGDGVRA-----VEITH-----GKNGHV 169
DB 595 VGVAAAGHGAVDQAAGRGDAPVLGDAAAVVALAVLLVEAGHGQPVAFQVRLEVGQGGAHA 654
QY 170 HV-----HALLMFGSDVSENILESPSDAMFD-----RMTSKLVS-LGFAAPL- 210
DB 555 GVGVAAGVAGEDVDHAL---AGDAGGVQVPAGAPVVLGIGGVAVEGLVADVGLUAATV 711
QY 211 -----RNSGGLDVRKIGTGEBADQVLAAYLTKTASGVGMVSGDGKSGRHRNAPW 260
DB 712 VDRPHGPGVQRLVGGLEIXRVRRAEQEAUVL-----VGVDVHLRCRVGH--- 756
QY 261 EIAVDVAG-GDPOALEM-----REFEFGSMGR-RATAWRSGLR-ARAG 301
DB 757 PVHADAVARGAEGAADVAGVIVAVHRAADAERRAVIDGAAGRHLVARGGGIRMARVE 816
QY 302 LGAEITDAQIIVEQESAPVMAIIPARSWMMIRTCAPVFG 342
DB 817 TGVHLADLHALVADGAGGILVGPFPFQAPVLL-----VFG 851

RESULT 4

US-09-252-991A-30345
; Sequence 30345, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30345
; LENGTH: 1395
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30345

Query Match 5.0%; Score 101; DB 4; Length 1395;
Best Local Similarity 24.5%; Pred. No. 0.57;
Matches 104; Conservative 34; Mismatches 162; Indels 124; Gaps 24;

QY 13 DRPVLVSSDKRGIRHRLRPKLQITTTSETNACOR--PISG-----VNGVTIVNGPKSGG 66
DB 511 DRIAVAGVEDHRAAR--FRSAEAAAQVQDGRGRRAIDGGAALAGVARRIGRGVD 568
QY 67 FGGLRS-CGKGCWICPCCAKGVGAHRADEISQVV-----AHQIG-----TGSVAMVTMTMRH 116

Db 569 HGSIRQRRGRG-EAPVAAA-VGGDLADRVAVAIGQGHAAARLGGAAERGAAGVGD-----H 622
QY 117 TAGO--RLHDLWTGLSAWKAATNG-----RRWTEREMVGC-----DGYV 155
Db 623 WRGRGYRVH-----GDTAARTGAVGGVGRGVDHGAVGQRRGRSEAPVAAAIGGDLPGDIA 678
QY 156 RAVETHGKNGWHVHALLMFSGDVSENILESFSDAMFDRWTSKLVSLGFAAPLR----- 211
Db 679 VAIQCGHG-----GARLGGAAERGAARVAVDHWGRCHGVHGDAGRRTGVA 723
QY 212 ----NSGGLDVKIG-----GEADQVLAAYLTKIASGVGMEVSGDGKSGRHNAPWEIA 263
Db 724 GRVSGRGVDHRAVGQRRARGEG-PVAAAAGVDLANRVAVAVGQHG-SARLGGAAEGRTV 781
QY 264 VDVGDPDQALELWREFFSGMGR-ALAW-----SRGLR-----ARAGLGA 305
Db 782 ARIDGGR-----RDGHNAGRRTAVAGGIPGHVHDHGTVRQGRGRERFPVAVGIGG 836
QY 306 LTDAQIVEQEEAPVMVAIIPARSMMIRTCAPVVFGEIILGLVEAGATMENLRDLHLVRL 365
Db 837 LAD-----RLAIVGQGHAAARFGSPAFAVARL-----DHRVR 873
QY 366 PAAD 369
Db 874 HAVD 877

RESULT 5

US-09-252-991A-23181
; Sequence 23181, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23181
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23181

Query Match 5.0%; Score 99.5; DB 4; Length 676;
Best Local Similarity 25.6%; Pred. No. 0.28;
Matches 89; Conservative 31; Mismatches 135; Indels 93; Gaps 19;
QY 62 PKSGFGGLRSCGKWTIC--PCCAGKVGARADEIS-----QVVAHQLTGSGVAMVTMT 113
Db 318 PRGPGAAGRGAGSGCGCGCGAGRTAGGASPVATGSGDCCSRAPAPVA----- 372
QY 114 MRHTAQORLHDLWTGLSAWKAATNGRMRTEREMYGCDGYVRVAVRITHGKNGWHVHVA 173
Db 373 -GRGAGNRRRRRWSLPAS--CLSRGR-----GVCRSVARRPG----- 409
QY 174 LMFSGDVSENILESFSDAMFDRWTSKLVSLGFAAPLRNSGGLDVKIGGEADQVLAAYL 233
Db 410 -AVRGSVSCALLTOAEKAAGD-----AVDLGLGA-----YRLDAQRFGENARQFLAACR 458
QY 234 TKIASGVG-MEVGS-----GDGK-----SGRHNAPWEIAYD--AVGSD-POALELWRE 279
Db 459 GKERRAFGETEAGLFQAHGEGRQLRDLVRQAGPAQFGFVVDHQAGGVDVAQLE----- 514
QY 280 FEPGSGMR--RAJAWSRGLRARAGLGAELTDAQIVEQEEAPV--MVAIIPARSMMIR 334
Db 515 -----GRMVAQTEQTEGVAE-----QGDVGEAQVQAQRGPAHRSVAVVDDAQSHLLA 564

QY 335 TCAPYV-----FGEILGLVEAGA--TWNELRDHLHLRPAAD 369
Db 565 GRAALVBLEEDHVPPRHDEGQQQAAGAAELAEERQAEHCHQHQDQAAD 612

RESULT 6

US-09-252-991A-19231
; Sequence 19231, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19231
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19231

Query Match 4.9%; Score 98.5; DB 4; Length 621;
Best Local Similarity 22.4%; Pred. No. 0.31;
Matches 62; Conservative 27; Mismatches 87; Indels 101; Gaps 11;
QY 175 LMFSGDVSENILESFSDAMFDRW-----TSKLVSLGFAAPLRNSGGL 216
Db 323 LLVSGGVDSNLLGYLDPQLQRHLCLEGEDEESLLPHRQLQRFELRQEAFFMLLRVAVG- 381
QY 217 DVKIGGEADQVLAAYLTKIASGVGME-----VCSGDKSGRHNAPWEIAYDVGDDPQ 272
Db 382 ---NFGATRMSSLLMYQLRADGIGEGYHCYVLLGEG-----ADELFWGYPR 425
QY 273 ALELWREFFSGMGRRAIARAGLGAELTDAQIVEQ-EESA----- 318
Db 426 HLELWRRRDAPEPRPFAAAWFGYRKAALLAEPAGRVARIEELAEALGQGLEAAIG 485
QY 319 --PVMVALIP-----ARSWMMIRTC--APYVFG-----EILG- 349
Db 486 QFDLHYSLEPLLRADHLLMSRTIEARTPYLHGALAQAGRLQRIVGDTAKAPIVALLEQ 545
QY 350 AGATWE-----NLRDHLHYRLPA 367
Db 546 REKRWQAQPKRHFRPLPFRMPQALGEMRRHLAERLPA 582

RESULT 7

US-09-252-991A-22282
; Sequence 22282, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22282
; LENGTH: 1409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22282	
Query Match 4.9%; Score 98.5; DB 4; Length 1409;	
Best Local Similarity 23.8%; Pred. No. 1;	
Matches 103; Conservative 51; Mismatches 160; Indels 119; Gaps 24;	
Qy	14 RPPVLSSDKGRTRHEL--PKLQIITSETFNACGRPISGVNGVTI-----VNGP 62
Db	275 RPFALDQSGGILQVQRCPGLQLSR-----RLASRLAGPAGETVDRPRDPLHLGLGY 330
Qy	63 KSGS-----FGGLRSCGKWIC--PCCAG-----KVGARADEISQVVAHQ 101
Db	331 RQSGADHPVAGALLBEIEGQLQAAIPEVVAAADPHLAALQORQAGARAETIQCAQLQ 390
Qy	102 L-----GT-----GSVAMVTWTRHTAGOR---LHDLWTGLSAANKAATN----- 138
Db	391 LAIAGLRQPGTGVQDVAAATAQSLAVGQQRPRGHVEQTALSOAQVAAAAQGNPSAVAV 450
Qy	139 GRR---WTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESPSDAMPD 195
Db	451 GQRIVLRQQQVAPQAAGIOPVAVAHQAQA-----VFAGIDETI-----DRPVR 496
Qy	196 RWTSKLVSLGFAAPLRNSSLGDLVRKIGGEADQVLAAYLTKIASGVGMVSGDGKSGRHG 255
Db	497 RNLQHVAGIQAAMGGDGVVD--RHVVGEQD--CLAAI-----GQQRADVADALAGRHL 548
Qy	256 NRAPWEIADVAGGDPQ---ALELWREFEFGSMGRRRAIAWSRGLRPARAGLGAELTDAQIV 312
Db	549 QAAQLE-AVEQFAVQAQAALAGGHRSVQGRLLGRRLAGTDHLLAAGIGAE-----SVL 603
Qy	313 EOBESAPVMVAIIP-----ARSMWMTTCAPYVFGELGIVEAGATWE-----NLR 358
Db	604 QQAALHVGVGAVPDDQLCAARHGIL-----RLVAVQAG-LWRFGLPAAVGVA 652
Qy	359 DHL--HYRLPAAD 369
Db	653 DHLAVEPRIGNAD 665
RESULT 8	
US-09-252-991A-30935	
; Sequence 30935, Application US/09252991A	
; Patent No. 6551795	
; GENERAL INFORMATION:	
; APPLICANT: Marc J. Rubenfield et al.	
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS	
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS	
; FILE REFERENCE: 107196.136	
; CURRENT APPLICATION NUMBER: US/09/252,991A	
; CURRENT FILING DATE: 1999-02-18	
; PRIOR APPLICATION NUMBER: US 60/074,788	
; PRIOR FILING DATE: 1998-02-18	
; PRIOR APPLICATION NUMBER: US 60/094,190	
; PRIOR FILING DATE: 1998-07-27	
; NUMBER OF SEQ ID NOS: 33142	
; SEQ ID NO 30935	
; LENGTH: 1245	
; TYPE: PRT	
; ORGANISM: Pseudomonas aeruginosa	
US-09-252-991A-30935	
Query Match 4.9%; Score 97.5; DB 4; Length 1245;	
Best Local Similarity 21.2%; Pred. No. 1.1;	
Matches 86; Conservative 50; Mismatches 143; Indels 127; Gaps 19;	
Qy	9 LSGKDRP--PVLVSSDKGRIRHELPRKQLQIITSETFNACGRPI-----SGVNGVT 57
Db	198 LRGQPSGPAFPGRQRRPGQPHRTYVQ-----RPAQRAGPRRGETGARGVA 235
Qy	58 IVNGPKSGFGGLRSCGKW-----ICPCCAGKVAHRADEISQVVAHQ 102
Db	236 AVPGE-----GCRTCAAGWRQAPGPRRAAPGTRPCGAGAGQLARWATGVPRQAAGRT 289
US-09-335-409-11	
; Sequence 11, Application US/09335409	
; Patent No. 6121029	
; GENERAL INFORMATION:	
; APPLICANT: Schupp, Thomas	
; APPLICANT: Ligon, James	
; APPLICANT: Molnar, Istvan	
; APPLICANT: Zirkle, Ross	
; APPLICANT: Cyr, Devon	
; APPLICANT: Goerlach, Joern	
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES	
; FILE REFERENCE: 4-30582A	
; CURRENT APPLICATION NUMBER: US/09/335,409	
; CURRENT FILING DATE: 1999-06-17	
; NUMBER OF SEQ ID NOS: 30	
; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 11	
; LENGTH: 713	
; TYPE: PRT	
; ORGANISM: Sorangium cellulosum	
US-09-335-409-11	
Query Match 4.8%; Score 97; DB 3; Length 713;	
Best Local Similarity 22.1%; Pred. No. 0.55; Indels 80; Gaps 12;	
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;	
Qy	30 LRPKLQIITSETFNACGRPISGVNGVTIIVNGPKSGFGGLRSCGKWICPCACGKVGAAH 89
Db	302 MRVDVSQLRTPAAGTVALLLATATAAKVVPAAALGARLGLRGSEAAALVAVLWKG-- 359
Qy	90 RADEISQVVAHQIGTGS-----VAMVTWTRHTAGORLHDLWTGLSAANKAATN 140
Db	360 -TDLIVAIVGVVELGLLSNEAYTMVAVVALTV---TASPALL-IMLEKRA---PPTQEE 410
Qy	141 RWRTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESPSDAMPDRWTSK 200
Db	411 SARLEREEARRAYIPGE---RILVPIVAHUPGFATDIVESIVAS----- 454
Qy	201 LVSLGFAAPLRNSSLGDLVRKIGGEADQVLAAYLTKIASGVGMVSGDGKSGRHN-- 257
Db	455 -----KRKLGETVD--ITELSVQQAQPPSPRAAGEASRGLARLARLV 496
Qy	258 APWEIADVAGDPOALELWREFEFGSMGRRRAIAWSRGLRAGLGAELTDAQIVQEES 317
Db	497 GIMQRRELRSIQAILRASRDHLLVIGARSPARAGMSF-----GRLQDAIVQRAESN 551
Qy	318 APVMVAIIPARSWMIRTCAPIYVFGELGIVEAGATWENLHLYRLPAADV 370
Db	552 VLVVVGDPAAE---RASARRILVPIG-----LEYSFAAADL 586

```
RESULT 10
US-09-568-102-11
; Sequence 11, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

QY 30 LRPKLOQITTSFTNACGRPISGVGTIYVNGPKSGFGGLRSCGKWCIPCCAGKVGAAH 89
Db 302 MRVDVSQRLTPAAWGTVALLLATATAAKVVPAAALGARLGLRGSEAAALVAVGLNMKG-- 359
QY 90 RADEISQVVAHQLTGS-----VAMVTMTMRHTAGORLHDLWTGLSAAMKAATNGR 140
Db 360 -TDLIVAIVGVVGLLSNEAYTMVAVVALTV-----TASPALL-IWLEKRA---PPTQEE 410
QY 141 RWRETERMYGCDGYVRAVEITHGKNGWHVHALLMPSGDVSENILESFSDFAMFDRWTSK 200
Db 411 SARLEREEAARRAYIPGVE-----RILVPIVAHALPGFATDIVESIVAS----- 454
QY 201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYITKIASGVGMVSGDGKSGRHN--- 257
Db 455 -----KRKLGETVD-ITELSEVQAPGPSRAAGEASRGLARLGLRV 496
QY 258 APWEIADVAGDPQALELWREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVEQES 317
Db 497 GIWRQRRELGRSGTQAILRASRDHLLVIGARSPARAGMSF-----GRLQDAIVQRAESN 551
QY 318 APWVAIIIPARSMWMTITCAPYVFGELGLVEAGATWENLRDLHLYRLPAADV 370
Db 552 VLWVVGDPPEAAE---RASARRILVPIIG-----LEYSFAAADL 586

RESULT 11
US-09-568-480-11
; Sequence 11, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-480-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

QY 30 LRPKLOQITTSFTNACGRPISGVGTIYVNGPKSGFGGLRSCGKWCIPCCAGKVGAAH 89
Db 302 MRVDVSQRLTPAAWGTVALLLATATAAKVVPAAALGARLGLRGSEAAALVAVGLNMKG-- 359
QY 90 RADEISQVVAHQLTGS-----VAMVTMTMRHTAGORLHDLWTGLSAAMKAATNGR 140
Db 360 -TDLIVAIVGVVGLLSNEAYTMVAVVALTV-----TASPALL-IWLEKRA---PPTQEE 410
QY 141 RWRETERMYGCDGYVRAVEITHGKNGWHVHALLMPSGDVSENILESFSDFAMFDRWTSK 200
Db 411 SARLEREEAARRAYIPGVE-----RILVPIVAHALPGFATDIVESIVAS----- 454
QY 201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYITKIASGVGMVSGDGKSGRHN--- 257
Db 455 -----KRKLGETVD-ITELSEVQAPGPSRAAGEASRGLARLGLRV 496
QY 258 APWEIADVAGDPQALELWREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVEQES 317
Db 497 GIWRQRRELGRSGTQAILRASRDHLLVIGARSPARAGMSF-----GRLQDAIVQRAESN 551
QY 318 APWVAIIIPARSMWMTITCAPYVFGELGLVEAGATWENLRDLHLYRLPAADV 370
Db 552 VLWVVGDPPEAAE---RASARRILVPIIG-----LEYSFAAADL 586

RESULT 11
US-09-567-969-11
; Sequence 11, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
```

```

Db      411 SARLREERAAARYIPGVE---RILVPTVAHALPGFATDIVESIVAS-----454
Qy      201 LVSLGFAAPLRNSGGLDKVRIKIGGEADOVLAAYLTKIASGVMEVSGDGKSGRHGNR---257
Db      455 -----KKKLGETVD-ITELSV EQQAPGPSRAAGEASRGLARGLARLV 496
Qy      258 APWEIADVAGGDPQALELWRREFEGSMGERRAIAMSRGLRAGLAGAELTDAQIVEQES 317
Db      497 GIWRQRRELRGSIQAILRASRDHLLVIGARSFARAGMSF-----GRLQDAIVQRAESN 551
Qy      318 APVMAIIPARSMWMTTCAPYVFGELGLVEAGATWENLRDHLHYRLPAADV 370
Db      552 VLVVVGDPPEAAE-----RASARRILVPIIG-----LEYSFAADL 586

RESULT 13
US-09-568-486-11
; Sequence 11, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

Qy      30  LREKLOQITSETNFNACGRPISGVNGVTIYVNGPKSGFGGLRSCGKGIWCPCCAGKVGAAH 89
Db      302 MRVDVQLRTPAAGTVALLLHATAAKVPAALGARLGLRGEAALVAVGLNMKG-- 359
Qy      90  RADEISQVVAHOLGTGS-----VAMVTMTMHTTAGORLHDLTGLSAAKAAATNGR 140
Db      360 -TDLIIVAIVGVELGLLSNEAYTWYAVVALTV---TASPALL-IMLEKRA---PPTQEE 410
Qy      141 RWRTEREMYGCDYVRAVEITHGKNGWHVHALLMPSGDVSNILEFSFSDAMFDRWTSK 200
Db      411 SAELEREERAAARYIPGVE---RILVPIVAHALPGFATDIVESIVAS-----454
Qy      201 LVSLGFAAPLRNSGGLDKVRIKIGGEADOVLAAYLTKIASGVMEVSGDGKSGRHGNR---257
Db      455 -----KKKLGETVD-ITELSV EQQAPGPSRAAGEASRGLARGLARLV 496
Qy      258 APWEIADVAGGDPQALELWRREFEGSMGERRAIAMSRGLRAGLAGAELTDAQIVEQES 317
Db      497 GIWRQRRELRGSIQAILRASRDHLLVIGARSFARAGMSF-----GRLQDAIVQRAESN 551
Qy      318 APVMAIIPARSMWMTTCAPYVFGELGLVEAGATWENLRDHLHYRLPAADV 370
Db      552 VLVVVGDPPEAAE-----RASARRILVPIIG-----LEYSFAADL 586

RESULT 14
US-09-568-472-11
; Sequence 11, Application US/09568472

```


; ORGANISM: Sorangium cellulosum
US-09-567-899-11

Query Match	4.8%;	Score 97;	DB 3;	Length 713;
Best Local Similarity	22.1%;	Pred. No. 0.55;		
Matches	78;	Conservative 38;	Mismatches 157;	Indels 80; Gaps 12;

QY	30	LRPKLOQIITTSFNFNACGRPI	SGVNGVTI	NGPKSGFG	GLRSCG	KWITCPC	CAGKVGAH	89
DB	302	MRVDVSQRLTPAAWGTIVALLI	ATAAKVPA	ALGARLGLR	GEAALVAV	GLMVGK	---	359
QY	90	RADEISQVVAHQIQTGS	-----	VAMVTMT	TRITAGOR	LHDLTGL	SAAKAATNGR	140
DB	360	-TDLIIVAI	GVGELGLLSNEAY	TYAVVALTV	-----	TASPALL	-IMLEKPA	410
QY	141	RWRETEREMYGCDYVRAVEI	THGKNGH	VHVHALLM	PSGDV	SNILLES	PSDAMFDRWTSK	200
DB	411	SARLEREEAARRAYIPGVE	-----	RILVPI	VAHALPG	FATDIVES	IVAS	454
QY	201	LVSGLGFAAPLRNSGGLDV	RKIGCEAD	QVLAAYLT	KTIAS	GVMEVGS	GDKSGRHN	257
DB	455	-----	KRKIGETVD	-ITEL	SVGEQAPG	PSRAAGE	BSRGLASGLARLRV	496
QY	258	APWEIADVAGGDPQAL	ELWREFE	PGSGMRRAL	AWSRGLR	ARAGLCAE	LTDAGIVQEES	317
DB	497	GIWRQRRRLRGSIQAIL	RASRDH	LLVIGARS	PARAGMSF	-----	GRQDAIVQRAESN	551
QY	318	APWVAIIIPARSMMIT	TCAPYV	FGIILG	VEAGATW	ENRDLH	HYRLPAADV	370
DB	552	VLVVGGDPFAAE	-----	RASARRIL	VIPIIG	-----	LEYSFAAADL	586

Search completed: October 23, 2004, 03:58:40
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 03:57:47 ; Search time 85 Seconds
(without alignments)
1443.583 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGSKDRPPVLVS.....HLHYRLPAADVRRPIISVRK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2005	100.0	379	14	US-10-007-527A-2
2	2005	100.0	379	14	US-10-007-452-2
3	2005	100.0	379	16	US-10-415-562A-2
4	609.5	30.4	459	14	US-10-007-527A-21
5	609.5	30.4	459	14	US-10-007-452-21
6	609.5	30.4	459	16	US-10-415-562A-21
7	326	16.3	427	9	US-09-835-381-6
8	322	16.1	451	14	US-10-007-527A-24
9	322	16.1	451	14	US-10-007-452-24
10	322	16.1	451	16	US-10-415-562A-24
11	318	15.9	427	9	US-09-835-381-8
12	318	15.9	427	14	US-10-196-232-7
13	316	15.8	456	14	US-10-007-527A-22

14	316	15.8	456	14	US-10-007-452-22	Sequence 22, Appl
15	316	15.8	456	16	US-10-415-562A-22	Sequence 22, Appl
16	282	14.1	466	9	US-09-826-191-9	Sequence 9, Appl
17	282	14.1	466	14	US-10-263-666-9	Sequence 9, Appl
18	282	14.1	466	14	US-10-261-481-9	Sequence 9, Appl
19	282	14.1	466	14	US-10-261-942-9	Sequence 9, Appl
20	266.5	13.3	528	14	US-10-007-527A-23	Sequence 23, Appl
21	266.5	13.3	528	14	US-10-007-452-23	Sequence 23, Appl
22	266.5	13.3	528	16	US-10-415-562A-23	Sequence 23, Appl
23	109.5	5.5	409	15	US-10-425-114-65990	Sequence 65990, A
24	103.5	5.2	417	14	US-10-156-761-10798	Sequence 10798, A
25	103	5.1	647	14	US-10-021-660-94	Sequence 94, Appl
26	102	5.1	907	16	US-10-437-963-156771	Sequence 156771, A
27	101	5.0	373	16	US-10-437-963-152208	Sequence 152208, A
28	101	5.0	823	16	US-10-437-963-156772	Sequence 156772, A
29	100.5	5.0	544	14	US-10-156-761-8540	Sequence 8540, Ap
30	99.5	5.0	468	14	US-10-369-493-12194	Sequence 12194, A
31	98.5	4.9	532	15	US-10-282-122A-43657	Sequence 43657, A
32	98.5	4.9	5835	15	US-10-378-083-20	Sequence 20, Appl
33	98	4.9	533	16	US-10-437-963-149861	Sequence 149861, A
34	97.5	4.9	515	16	US-10-437-963-187892	Sequence 187892, A
35	97	4.8	559	14	US-10-369-493-4607	Sequence 4607, Ap
36	97	4.8	567	14	US-10-369-493-7364	Sequence 7364, Ap
37	97	4.8	713	13	US-10-014-717-11	Sequence 11, Appl
38	96.5	4.8	938	16	US-10-437-963-189869	Sequence 189869, A
39	96	4.8	7746	14	US-10-156-761-7965	Sequence 7965, Ap
40	95.5	4.8	3798	13	US-10-014-717-6	Sequence 6, Appl
41	95.5	4.8	4471	14	US-10-205-032-10	Sequence 10, Appl
42	95	4.7	184	16	US-10-437-963-181278	Sequence 181278, A
43	95	4.7	321	16	US-10-437-963-183714	Sequence 183714, A
44	94.5	4.7	552	14	US-10-369-493-15887	Sequence 15887, A
45	94.5	4.7	552	16	US-10-437-963-135632	Sequence 135632, A

ALIGNMENTS

RESULT 1
US-10-007-527A-2
; Sequence 2, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CIL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rhodococcus AN12
US-10-007-527A-2

Query Match 100.0%; Score 2005; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSVSAEHLGSKDRPPVLVSDDKRGIRHRLPKLQOITTSFNACGRPIISGVNGVTIVN 60
|||||

Db 1 MTSVSAEHLGSKDRPPVLVSDDKRGIRHRLPKLQOITTSFNACGRPIISGVNGVTIVN 60
|||||

QY 61 GPKGSFGGLRSCGKMWICPCACAGKGAHRADEISQVAFHQLGTGSAVMVMTMRHTAGQ 120
|||||

Db 61 GPKGSFGGLRSCGKMWICPCACAGKGAHRADEISQVAFHQLGTGSAVMVMTMRHTAGQ 120
|||||

QY 121 RLHDLWTGLSAWAATNGRWRTEREMYGCDGYVRAVEITHGKNGHVVHALLMFSGD 180
|||||

Db 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEQESAPVWVAIIPARSWMMIRTCPYVFGIILGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEQESAPVWVAIIPARSWMMIRTCPYVFGIILGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVPRPIISVRK 379
Db 361 LHVRLPAADVPRPIISVRK 379
RESULT 2
US-10-007-452-2
; Sequence 2, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rhodococcus AN12
US-10-007-452-2
Query Match 100.0%; Score 2005; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGSKDRPPVLVSSDKRGIRHELPRKLOQITTSFNAACGRPIISGVNGVTIVN 60
Db 1 MTSVSAEHLGSKDRPPVLVSSDKRGIRHELPRKLOQITTSFNAACGRPIISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTWHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTWHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEQESAPVWVAIIPARSWMMIRTCPYVFGIILGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEQESAPVWVAIIPARSWMMIRTCPYVFGIILGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVPRPIISVRK 379
Db 361 LHVRLPAADVPRPIISVRK 379
RESULT 4
US-10-007-527A-21
; Sequence 21, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97

Db 361 LHVRLPAADVPRPIISVRK 379
RESULT 3
US-10-415-562A-2
; Sequence 2, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rhodococcus AN12
US-10-415-562A-2
Query Match 100.0%; Score 2005; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGSKDRPPVLVSSDKRGIRHELPRKLOQITTSFNAACGRPIISGVNGVTIVN 60
Db 1 MTSVSAEHLGSKDRPPVLVSSDKRGIRHELPRKLOQITTSFNAACGRPIISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTWHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTWHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEQESAPVWVAIIPARSWMMIRTCPYVFGIILGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEQESAPVWVAIIPARSWMMIRTCPYVFGIILGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVPRPIISVRK 379
Db 361 LHVRLPAADVPRPIISVRK 379
RESULT 4
US-10-007-527A-21
; Sequence 21, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97

```

; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Arcanobacterium pyogenes
US-10-007-527A-21

Query Match      30.4%; Score 609.5; DB 14; Length 459;
Best Local Similarity 36.3%; Pred. No. 4.5e-49;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAHLSGKDRPPVLVSSD-----KRGIRHELRLPKLQIITSETFNACGR-PISGVNGV 56
Db 67 SATHPLGNTVLTFFVSNESKKTAKRSRSEYELRDGLAEISTIESVRKCGRPVAPLVSL 126

QY 57 TIVNGPKGSFGGLRSCGKWI CPCAGKVGARADEISQVVAHQLTG-SVAMVTMTMR 115
Db 127 RAKSDGKAGYGGIHTCGSVWACPVCSAKIAARRKTDLQQVVDHAVKHGTMVSMILTQR 186

QY 116 HTAGORLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
Db 187 HHKQGLKHLWDALSTAWNRVTSGRRMTFEKQFGLGVYVRANEITHGKHGWHVSHVLI 246

QY 176 MFGSDVSENILESF-----SDAMFDRWTSKLVSLGFAAPLRNSGG 215
Db 247 ISEKD---PLTSTFVYQKQGRRLPYPPPEIYMSSDFFIAERWEAGLAKHGVD-F-LRDSGG 302

QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA 266
Db 303 LDWTAVKADARAIGN-----YVSKMOTSTDAISSEVTLGFFKKARNGNRTFPQILADI 354

QY 267 VG-GDQOALELWRPEFSGMGRRAIAWSRGLRPARAGLGAELTDAQIVEQESAPVMVAII 325
Db 355 LSLGVDVDDLKWKKEYEKASFRRALTWKGLRDWANLGVQESDEEIA-SEEIGDEAIALF 413

QY 326 PARSMWMLRTCAPYVFG--EILGLVEAGA-----TWNELRDHLHYRLP 366
Db 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAYRWLDFRE-IDWSLP 455

RESULT 5
US-10-007-452-21
; Sequence 21, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Arcanobacterium pyogenes
US-10-007-452-21

Query Match      30.4%; Score 609.5; DB 14; Length 459;
Best Local Similarity 36.3%; Pred. No. 4.5e-49;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAHLSGKDRPPVLVSSD-----KRGIRHELRLPKLQIITSETFNACGR-PISGVNGV 56
Db 67 SATHPLGNTVLTFFVSNESKKTAKRSRSEYELRDGLAEISTIESVRKCGRPVAPLVSL 126

QY 57 TIVNGPKGSFGGLRSCGKWI CPCAGKVGARADEISQVVAHQLTG-SVAMVTMTMR 115
Db 127 RAKSDGKAGYGGIHTCGSVWACPVCSAKIAARRKTDLQQVVDHAVKHGTMVSMILTQR 186

QY 116 HTAGORLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
Db 187 HHKQGLKHLWDALSTAWNRVTSGRRMTFEKQFGLGVYVRANEITHGKHGWHVSHVLI 246

QY 176 MFGSDVSENILESF-----SDAMFDRWTSKLVSLGFAAPLRNSGG 215
Db 247 ISEKD---PLTSTFVYQKQGRRLPYPPPEIYMSSDFFIAERWEAGLAKHGVD-F-LRDSGG 302

QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA 266
Db 303 LDWTAVKADARAIGN-----YVSKMOTSTDAISSEVTLGFFKKARNGNRTFPQILADI 354

QY 267 VG-GDQOALELWRPEFSGMGRRAIAWSRGLRPARAGLGAELTDAQIVEQESAPVMVAII 325
Db 355 LSLGVDVDDLKWKKEYEKASFRRALTWKGLRDWANLGVQESDEEIA-SEEIGDEAIALF 413

QY 326 PARSMWMLRTCAPYVFG--EILGLVEAGA-----TWNELRDHLHYRLP 366
Db 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAYRWLDFRE-IDWSLP 455

```

```

QY 116 HTAGORLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
Db 187 HHKQGLKHLWDALSTAWNRVTSGRRMTFEKQFGLGVYVRANEITHGKHGWHVSHVLI 246

QY 176 MFGSDVSENILESF-----SDAMFDRWTSKLVSLGFAAPLRNSGG 215
Db 247 ISEKD---PLTSTFVYQKQGRRLPYPPPEIYMSSDFFIAERWEAGLAKHGVD-F-LRDSGG 302

QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA 266
Db 303 LDWTAVKADARAIGN-----YVSKMOTSTDAISSEVTLGFFKKARNGNRTFPQILADI 354

QY 267 VG-GDQOALELWRPEFSGMGRRAIAWSRGLRPARAGLGAELTDAQIVEQESAPVMVAII 325
Db 355 LSLGVDVDDLKWKKEYEKASFRRALTWKGLRDWANLGVQESDEEIA-SEEIGDEAIALF 413

QY 326 PARSMWMLRTCAPYVFG--EILGLVEAGA-----TWNELRDHLHYRLP 366
Db 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAYRWLDFRE-IDWSLP 455

RESULT 6
US-10-415-562A-21
; Sequence 21, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Arcanobacterium pyogenes
US-10-415-562A-21

Query Match      30.4%; Score 609.5; DB 16; Length 459;
Best Local Similarity 36.3%; Pred. No. 4.5e-49;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAHLSGKDRPPVLVSSD-----KRGIRHELRLPKLQIITSETFNACGR-PISGVNGV 56
Db 67 SATHPLGNTVLTFFVSNESKKTAKRSRSEYELRDGLAEISTIESVRKCGRPVAPLVSL 126

QY 57 TIVNGPKGSFGGLRSCGKWI CPCAGKVGARADEISQVVAHQLTG-SVAMVTMTMR 115
Db 127 RAKSDGKAGYGGIHTCGSVWACPVCSAKIAARRKTDLQQVVDHAVKHGTMVSMILTQR 186

QY 116 HTAGORLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
Db 187 HHKQGLKHLWDALSTAWNRVTSGRRMTFEKQFGLGVYVRANEITHGKHGWHVSHVLI 246

QY 176 MFGSDVSENILESF-----SDAMFDRWTSKLVSLGFAAPLRNSGG 215
Db 247 ISEKD---PLTSTFVYQKQGRRLPYPPPEIYMSSDFFIAERWEAGLAKHGVD-F-LRDSGG 302

QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA 266
Db 303 LDWTAVKADARAIGN-----YVSKMOTSTDAISSEVTLGFFKKARNGNRTFPQILADI 354

QY 267 VG-GDQOALELWRPEFSGMGRRAIAWSRGLRPARAGLGAELTDAQIVEQESAPVMVAII 325
Db 355 LSLGVDVDDLKWKKEYEKASFRRALTWKGLRDWANLGVQESDEEIA-SEEIGDEAIALF 413

QY 326 PARSMWMLRTCAPYVFG--EILGLVEAGA-----TWNELRDHLHYRLP 366
Db 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAYRWLDFRE-IDWSLP 455

```

```

RESULT 7
US-09-835-381-6
; Sequence 6, Application US/09835381
; Patent No. US20020045223A1
; GENERAL INFORMATION:
; APPLICANT: SUGA, Mikiko
; APPLICANT: ASAKURA, Yoko
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hideo
; APPLICANT: KURASHI, Osamu
; TITLE OF INVENTION: ARGININE REPRESSOR DEFICIENT STRAIN OF CORYNEFORM BACTERIA
; TITLE OF INVENTION: FOR PRODUCING L-ARGININE
; FILE REFERENCE: 206018USO
; CURRENT APPLICATION NUMBER: US/09/835,381
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: JP2000-129167
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-835-381-6

Query Match 16.3%; Score 326; DB 9; Length 427;
Best Local Similarity 29.1%; Pred. No. 3.7e-22;
Matches 98; Conservative 58; Mismatches 131; Indels 50; Gaps 15;

Qy 34 LQIITSETENACGR-----PISGVNGVTIVNGPKSGFGGLRSCGKGWICPCCAKV 86
Db 27 MYKITNSKALAGCHWRDRDEAVVSWSS-----NG--ASQEGQLNHSRHWGSLAELEV 79

Qy 87 GAHRADBSIQVAHQGLGTGSVAMVTM-TMRHTAQRLHDLMTGLSAAWKAATNGRRMTE 145
Db 80 MGERRIELATATKNHLAAGGALMMFVGTVRHNRSQSFAQVEAGIKTAYSSMVKTSQWKKE 139

Qy 146 REMVCGDGYRVAEITHG-KNGMHVHVHALLMFSODUSENTILEFSFSDAMFDRWTSKLYSL 204
Db 140 RARYGVHTYSYDEVETDSWANGHNLHRLNMLFLDRPLSDDELKAFEDSMFWSAGVKA 199

Qy 205 GFAPAPLRNSG-GLD-VRKIGEADQVLAAYITKTIASGVGMVEGSDGSGRHNRPWEI 262
Db 200 GMDAPLREHGVKLDQVSTWGGDAK-MATYLAK---GMSQLTGSATKTSKGSYTFQM 255

Qy 263 AVDAVGDPQALE-----LWREFSGMGRRAIAMSRLGARAGLGAELTDAQI--- 311
Db 256 -LDMLAQSDAGEDMDAVLVARVREYEGVKNLRS-SWSRG--AKRALGDIYDADVRRE 311

Qy 312 VEQS-----ESAPVWVAIIIPARSMWMIRT 335
Db 312 MEELYKLAGLEAPERVESTRVAVLVKPDWKLIQS 348

RESULT 8
US-10-007-527A-24
; Sequence 24, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97

```

```

QY 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGGLDVRKIGGEAD-QVLAAYLTAK 237
Db 175 EPSEALDEWQGWRAV-WTAALKVNPQTPDDRHGVDFKRLETERDANDLAETAKQ 233
QY 238 SG--VGMVEVSGDGKSGRGNRAPWEI---AVDAVGG---DPOA-----LELWREPE 281
Db 234 DGKAPALELARADLKTANGGNVAPFELLGRIGDLTGMTEDDAAGVGSLEWNLARWHEYE 293
QY 282 FGSWGRRAIWSGLRARGAGLGAELTDA 309
Db 294 RATKGRRALEWTRYLRQMLGLDGGDTEA 321

RESULT 10
US-10-415-562A-24
; Sequence 24, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CLI709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; PRIORITY FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Streptomyces nigirificiens
US-10-415-562A-24

Query Match 16.1%; Score 322; DB 16; Length 451;
Best Local Similarity 30.8%; Pred. No. 9.4e-22;
Matches 101; Conservative 41; Mismatches 110; Indels 76; Gaps 15;

QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKWCIPCCAGKVGAAHRADEISQVVAHOL 102
Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRILWCLPVCATIRHKRAEETAAVVEWI 54
QY 103 GTGSVA-MVTMTMHTAGORLHDLMTGLSAANKAATNGRR-----WRTE---- 145
Db 55 KRGGTAVLVFTFARHGHTDRLADLMDALQGTRKTADAPRRPGAYORLITGTWAGRAKD 114
QY 146 -----REMYGCDGYVRAVEITHGK-NGWHVHVHALLMFSGDV-----SENILESF 189
Db 115 GHRAADREGIRDRTGYVGMIRATEVTVGQINGMHPHIAIHLVGGRTGERSAKQIVGTF 174
QY 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGGLDVRKIGGEAD-QVLAAYLTAK 237
Db 175 EPSEALDEWQGWRAV-WTAALKVNPQTPDDRHGVDFKRLETERDANDLAETAKQ 233
QY 238 SG--VGMVEVSGDGKSGRGNRAPWEI---AVDAVGG---DPOA-----LELWREPE 281
Db 234 DGKAPALELARADLKTANGGNVAPFELLGRIGDLTGMTEDDAAGVGSLEWNLARWHEYE 293
QY 282 FGSWGRRAIWSGLRARGAGLGAELTDA 309
Db 294 RATKGRRALEWTRYLRQMLGLDGGDTEA 321

```

```

RESULT 11
US-09-835-381-8
; Sequence 8, Application US/09835381
; Patent No. US20020045223A1
; GENERAL INFORMATION:
; APPLICANT: SUGA, Mikiko
; APPLICANT: ASAKURA, Yoko
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; APPLICANT: KURAHASHI, Osamu

; TITLE OF INVENTION: ARGININE REPRESSOR DEFICIENT STRAIN OF CORYNEFORM BACTERIUM AND
; FILE REFERENCE: 206018US0
; CURRENT APPLICATION NUMBER: US/09/835,381
; PRIOR FILING DATE: 2001-04-17
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-835-381-8

Query Match 15.9%; Score 318; DB 9; Length 427;
Best Local Similarity 28.8%; Pred. No. 2.1e-21;
Matches 97; Conservative 58; Mismatches 132; Indels 50; Gaps 15;

QY 34 LQOITSETFNACGR-----PIGVNGVTIVNGPKSGFGGLRSCGKWCIPCCAGKV 86
Db 27 MYKITNSKALAGCHWRERDEAVVSWSS-----NG--ASQFEGLONSHRWSSSLAELEV 79
QY 87 GAHRADEISQVVAHOLGTGSVAMVTM-TMRHTAGORLHDLMTGLSAANKAATNGRRWRTE 145
Db 80 MGERRIELAIATKNHLAAGGALMMFVGTVRHNRSSQFAQVEAGIKTAYSSMVVKTSQWKKE 139
QY 146 REMYGCDGYVRAVEITHG-KNGWHVHVHALLMFSGDVSENILESFSDAMFDRWTSKLVSL 204
Db 140 RARYGVEHTYSDYEVTSWANGWHLHRNMLFLDRLPSDDDELKAFEDSMFSRWSAGVYKA 199
QY 205 GFAAPLRNSG-GLD-VRKIGGEADQVLAAYLTAKIASGVGMEVSGDGKSGRGNRAPWEI 262
Db 200 GMDAPLRHGVKLDQVSTWGGDAK-MATYLAK---GMSQELTGSATKTSKSGSYTFQM 255
QY 263 AVDAVGGDPQALE-----LWREFEFGSMGRRAIWSGLRARGAGLGAELTDAQI--- 311
Db 256 -LMDLADQSDAGEDMDAVLVARWREYEVGSKNLR-SWSRG--AKRALGIDYIDADVRE 311
QY 312 VEQD-----ESAPVVMVAIIPARSWMMIRT 335
Db 312 MEELYKLGLAEPERVESTRVAVALKVDPDDWKLIQS 348

RESULT 12
US-10-196-232-7
; Sequence 7, Application US/10196232
; Publication No. US20030113899A1
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, MIKIKO
; APPLICANT: ITO, HISAO
; APPLICANT: GUNJI, YOSHIVA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 225391US0
; CURRENT APPLICATION NUMBER: US/10/196,232
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: JP 2001-224586
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-10-196-232-7

Query Match 15.9%; Score 318; DB 14; Length 427;
Best Local Similarity 28.8%; Pred. No. 2.1e-21;
Matches 97; Conservative 58; Mismatches 132; Indels 50; Gaps 15;

QY 34 LQOITSETFNACGR-----PIGVNGVTIVNGPKSGFGGLRSCGKWCIPCCAGKV 86

```


Db 27 MYKITNSKALAGCHRRERDRDAVAVSWSS-----NG--ASQFGLQNSHRSWGSLSAELEV 79

Qy 87 GAHRADBSIQVVAHQLTGVSAMVTM--TMRHTAGQRLHDLWTGLSAAWKAATNGRRWRTE 145

Db 80 MGERREIATATXNHLAAGALMMFVCTVHRNSQSFAQVEAGIKTAYSSWKTQWKE 139

Qy 146 REMYGCDDYVRAVEITHG-KNGWHVHVHALLMFGSDVSENLISFSDAMFDRWTSKLVS 204

Db 140 RARYGVHTYSDYEVTSWANGWHLHRNMLFLDRPLSDDELKAFEDSMFSRWSAGVYKA 199

Qy 205 GFAPLRNSG-GLD-VRKIGERADQVLAAYLTKTASGVGMEVSGDGKSGRHNRAPEWI 262

Db 200 GMDAPREHGVKLDQVSTWGGDAK-WATYLA-----GMSQLTSATKTASKSYTFQM 255

Qy 263 AVDVGDDPOALE-----LWREFFGSMGRRAIAWSRGLRARAGLGAELTDAQI--- 311

Db 256 -LMDAQSDAGEDMDAVLVARWEYEVGSKNLS-SWSRG--AKRALGIDYIDADVRE 311

Qy 312 VEQE-----ESAPVMAIIPARSMWIRT 335

Db 312 MEELYKLAGLEAPERVESTRAVALVKPDDWKLQIS 348

RESULT 13

US-10-007-527A-22

Sequence 22, Application US/10007527A

Publication No. US20030044807A1

GENERAL INFORMATION:

APPLICANT: Tomb, Jean-Francois

APPLICANT: Bramucci, Michael G.

APPLICANT: Cheng, Qiong

APPLICANT: Kostichka, Kristy N.

TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

FILE REFERENCE: CL1709 US NA

CURRENT APPLICATION NUMBER: US/10/007,527A

CURRENT FILING DATE: 2001-12-05

PRIOR APPLICATION NUMBER: 60/254,868

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Microsoft Office 97

SEQ ID NO 22

LENGTH: 456

TYPE: PRT

ORGANISM: Streptomyces lividans

US-10-007-527A-22

Query Match 15.8%; Score 316; DB 14; Length 456;

Best Local Similarity 30.5%; Pred. No. 3.6e-21;

Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

Qy 49 PISGV-----NGVTIVNGPKSGFGGLRSCKGKWTCCPCAGKVGHAHRADEISQVVAHQ 102

Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRIMLCPVCAATIRHKRAEITAAVVEWI 54

Qy 103 GTGSVA-MVTMTMRHTAGORLHDLWTGLSAAWKAATNGRR-----WRTE---- 145

Db 55 KRGGTAYLVTFTRARHGHTDLADLMDALQGTTRTPDSPRRPGAYORLITGGTWAGRAKD 114

Qy 146 -----REMYGCDGVYRAVEITHGK-NGWHVHVHALLMFGSDV-----SENILESF 189

Db 115 GHRAADREGIRDRIYGVGMIRATEVTVGQINGWHPHIAVLVGGRTGERSAKQIVATF 174

Qy 190 --SDAMFDRWTSKLVSGLFAAPLR-----NSGGLDVRKIGGEAD-OVLAAYLTAKIA 237

Db 175 EPTGAALDEWQHWSRV-WTAALRKVNPAFTPDDRHGVDVFKLETERDANDLAEYIAKTQ 233

Qy 238 SG--VGMVEVSGDGKSGRHNRAPEWI---AVDAVGG---DPOA-----LELWREFE 281

Db 234 DGKAPALELARADLKTATGNGVAPPELLGRIGDLTGGMTEDDAAGVGSLEWNLRSWHEYE 293

Qy 282 FGSNGRRRAIAWSRGLRARAGLGAELTDA 309

Db 294 RATGRRRAIEWTRVLRQMLGDLGGDTFA 321

RESULT 14

US-10-007-452-22

Sequence 22, Application US/10007452

Publication No. US20030093701A1

GENERAL INFORMATION:

APPLICANT: Tomb, Jean-Francois

APPLICANT: Bramucci, Michael G.

APPLICANT: Cheng, Qiong

APPLICANT: Kostichka, Kristy N.

TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

FILE REFERENCE: CL1709 US NA

CURRENT APPLICATION NUMBER: US/10/007,452

CURRENT FILING DATE: 2001-11-08

PRIOR APPLICATION NUMBER: 60/254,868

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Microsoft Office 97

SEQ ID NO 22

LENGTH: 456

TYPE: PRT

ORGANISM: Streptomyces lividans

US-10-007-452-22

Query Match 15.8%; Score 316; DB 14; Length 456;

Best Local Similarity 30.5%; Pred. No. 3.6e-21;

Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

Qy 49 PISGV-----NGVTIVNGPKSGFGGLRSCKGKWTCCPCAGKVGHAHRADEISQVVAHQ 102

Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRIMLCPVCAATIRHKRAEITAAVVEWI 54

Qy 103 GTGSVA-MVTMTMRHTAGORLHDLWTGLSAAWKAATNGRR-----WRTE---- 145

Db 55 KRGGTAYLVTFTRARHGHTDLADLMDALQGTTRTPDSPRRPGAYORLITGGTWAGRAKD 114

Qy 146 -----REMYGCDGVYRAVEITHGK-NGWHVHVHALLMFGSDV-----SENILESF 189

Db 115 GHRAADREGIRDRIYGVGMIRATEVTVGQINGWHPHIAVLVGGRTGERSAKQIVATF 174

Qy 190 --SDAMFDRWTSKLVSGLFAAPLR-----NSGGLDVRKIGGEAD-OVLAAYLTAKIA 237

Db 175 EPTGAALDEWQHWSRV-WTAALRKVNPAFTPDDRHGVDVFKLETERDANDLAEYIAKTQ 233

Qy 238 SG--VGMVEVSGDGKSGRHNRAPEWI---AVDAVGG---DPOA-----LELWREFE 281

Db 234 DGKAPALELARADLKTATGNGVAPPELLGRIGDLTGGMTEDDAAGVGSLEWNLRSWHEYE 293

Qy 282 FGSNGRRRAIAWSRGLRARAGLGAELTDA 309

Db 294 RATGRRRAIEWTRVLRQMLGDLGGDTFA 321

RESULT 15

US-10-415-562A-22

Sequence 22, Application US/10415562A

Publication No. US20040115661A1

GENERAL INFORMATION:

APPLICANT: E.I. du Pont De Nemours and Company

TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

FILE REFERENCE: CL1709 US PCT

CURRENT APPLICATION NUMBER: US/10/415,562A

CURRENT FILING DATE: 2003-11-17

PRIOR APPLICATION NUMBER: 60/254,868

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Microsoft Office 97

SEQ ID NO 22

LENGTH: 456

TYPE: PRT

ORGANISM: Streptomyces lividans

US-10-415-562A-22

Search completed: October 23, 2004, 04:03:56
Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 03:58:47 ; Search time 5309 Seconds

(without alignments)
3375.927 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005

Sequence: 1 MTSVSAEHLGSKDRPPVLVS.....HLHYRLPAADVRPPIISVRK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO.spool/US10007527/runat_22102004.114116.26617/app_query.fasta_1.519
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10007527@cgn 1_1_3731_@runat_22102004.114116.26617 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2005	100.0	1140	6	AX548642
C 2	2005	100.0	6334	1	AY178757 Rhodococc
C 3	2005	100.0	6334	6	AX548646
C 4	2005	100.0	9652	6	AX548648 Sequence

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 5	2005	100.0	11241	6	AX548647
C 6	2005	100.0	11241	12	AY180162
C 7	609.5	30.4	2439	1	APU83788
C 8	569	28.4	3540	1	AY150274
C 9	497	24.8	2051	1	PF66662
C 10	453	22.6	2661	1	SLMPG33A
C 11	452.5	22.6	9367	1	SN243257
C 12	425.5	21.2	4603	1	AY172684
C 13	409	20.4	300717	1	AE017229
C 14	399.5	19.9	2297	6	E17316
C 15	392.5	19.6	5750	1	AF085719
C 16	374.5	18.7	5804	7	AF428260
C 17	360.5	18.0	12855	1	AF533985
C 18	346.5	17.3	10992	1	STMPN22
C 19	345	17.2	3748	12	AB032347
C 20	345	17.2	3784	12	AB032348
C 21	345	17.2	5128	1	AB084167
C 22	342.5	17.1	4109	1	CCA308231
C 23	337	16.8	5349	12	AY072037
C 24	337	16.8	6342	12	AY072040
C 25	337	16.8	6607	12	AY072039
C 26	337	16.8	7560	6	AR475430
C 27	337	16.8	8058	12	AY072038
C 28	337	16.8	8215	12	PI414669
C 29	337	16.8	8830	1	PIJ101CG
C 30	337	16.8	12738	6	AR475429
C 31	333.5	16.6	2297	1	AF273219
C 32	326	16.3	3147	12	AF092035
C 33	326	16.3	4447	1	AF092037
C 34	326	16.3	4447	6	BD106940
C 35	326	16.3	4447	6	BD106940
C 36	326	16.3	4447	6	BD139585
C 37	326	16.3	4457	1	BLPLASML
C 38	326	16.3	5135	12	SVE133193
C 39	326	16.3	5216	12	SVE133194
C 40	326	16.3	6241	12	AF092036
C 41	326	16.3	6272	12	AF091268
C 42	326	16.3	6398	12	AF092931
C 43	326	16.3	6601	12	SVE133195
C 44	326	16.3	8161	12	AY585307
C 45	318	15.9	4447	6	AR173115

ALIGNMENTS

RESULT 1	AX548642	Sequence 1 from Patent WO02055709.	1140 bp	DNA	linear	PAT 27-NOV-2002
AX548642	Sequence 1 from Patent WO02055709.					
LOCUS	AX548642					
DEFINITION	AX548642					
ACCESSION	AX548642					
VERSION	AX548642.1	GI:25813612				
KEYWORDS						
SOURCE	Rhodococcus erythropolis					
ORGANISM	Rhodococcus erythropolis					
REFERENCE	1					
AUTHORS	Bramucci,M.G., Cheng,Q., Kostichka,K.N. and Tomb,J.F.					
TITLE	Rhodococcus cloning and expression vectors					
JOURNAL	Patent: WO 02055709-A.1 18-JUL-2002;					
FEATURES	E.I. DU PONT DE NEMOURS AND COMPANY					
source	Location/Qualifiers					
	1. .1140					
	/organism="Rhodococcus erythropolis"					
	/mol_type="unassigned DNA"					
	/strain="ANI2"					
	/db_xref="taxon:1833"					

ORIGIN

Alignment Scores:					
Pred. No.:	6,77e-123	Length:	1140		
Score:	2005.00	Matches:	379		
Percent Similarity:	100.00%	Conservative:	0		

Best Local Similarity: 100.00%		Mismatches: 0	
Query Match: 100.00%		Indels: 0	
DB: 6		Gaps: 0	
US-10-007-527A-2 (1-379) x AX548642 (1-1140)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyValAspArgProProValLeuValSer	20
Db	1	ATGACCGAGTAAAGTGTGACACCTTTCCGGCAAGACCGGCTCCCGCTCGTGTG	60
Qy	21	SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer	40
Db	61	TCCGATAGCGCGCATCCGCGACCACTCGGACCCAACTTCAACAAATCACACGTC	120
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	121	GAACATTTAACGCTGTGGCGCGCATTTCTGGCGTGAACGGTGTGACCAITGTCAAC	180
Qy	61	GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro	80
Db	181	GGTCCGAAAGGTTCTGGATTTCGGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCC	240
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValAlaHis	100
Db	241	TGCTGTCCGGGAAAGTCGTGCACATCGTCGACAGCAATTTCTCAAGTTGTGCTCAT	300
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	301	CAACTCGGACTGTGATCTGTTCGATGTGACGATGACCATCGCCCATACAGCTGGTC	360
Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140
Db	361	CGGCTCCACGACCTATGACTTGGACTTTCGGCAGCCTGGAAGCTGCGACCAACGGTC	420
Qy	141	ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160
Db	421	CGTTGGCTAGCAAGCTGAATGTACGGCTGCGACGATACGTGCGGCTGTTGAAATC	480
Qy	161	ThrHisGlyLysAsnGlyTyrPheValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	481	ACTCACGAAAAAAGCGGTGGCGCTCCACGCTTCCGCGCTTACTCATGTTCAGTGGT	540
Qy	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys	200
Db	541	GTGAGTGAGAACATCTCGAATCTTTCGGATGCGATGTCGATCGGTGACTTCCAA	600
Qy	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220
Db	601	CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTCGGGTGGTCTCGATGTACG	660
Qy	221	IleGlyGlyGluAlaAspGlnValLeuAlaIleValLeuThrLysIleAlaSerGlyVal	240
Db	661	ATCGCGGTGAAGCTGATCAAGTTCCTCGCTCGTATCTGACGAAATTCATCTGGCG	720
Qy	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr	260
Db	721	GGTATGGAGGTTGGTAGTGGCGCAAGAAAGTGTGTGATGCGCAACCGTGACCC	780
Qy	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe	280
Db	781	GAATTCGCTGTGTGATGTCAGTGGCGGGGATCCACAAAGCTTGGAACTGTGGCG	840
Qy	281	GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla	300
Db	841	GAGTTTGGTTGATCGGACGTCGGGCAATTCGGTGGTCCCGGATTCGGTCCCGAG	900
Qy	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGlnGluSerAlaProVal	320
Db	901	GGTCTTGGGCGAAGACTACAGATGCTCAGATCGTTGTAGCAGCAAGATCTGCCCG	960
Qy	321	MetValAlaIleIleProAlaArgSerTyrMetMetIleArgThrCysAlaProTyrVal	340
Db	961	ATGGTTGCGATCATTTCCGCGCGGATCGTGGATGATGATTCGGACTTGTGCGCTT	1020

Qy	341	PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTyrGluAsnLeuArgAspHis	360
Db	1021	TTGCGCGAGATCTCGACTCTGTCGAAGCTGGCGGCAATCTTCGTTGATCAC	1080
Qy	361	LeuHisTyrArgLeuProAlaAlaAspValArgProIleIleSerValArgLys	379
Db	1081	TTGCATTCGATTGCCCGCAGCGATGTGCGCCCCCGATAATATATCGTTCCGCAAG	1137
RESULT 2			
AY178757/c			
LOCUS			
Rhodococcus erythropolis cryptic plasmid pAN12, complete sequence.			
ACCESSION			
AY178757			
VERSION			
AY178757.1 GI:28628256			
KEYWORDS			
SOURCE			
ORGANISM			
Rhodococcus erythropolis			
Rhodococcus erythropolis			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacterineae; Nocardiaceae; Rhodococcus.			
REFERENCE			
AUTHORS			
Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and			
Cheng, Q.			
TITLE			
A small cryptic plasmid from Rhodococcus erythropolis:			
characterization and utility for gene expression			
JOURNAL			
MEDLINE			
PUBMED			
22718480			
12835922			
REFERENCE			
AUTHORS			
Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and			
Cheng, Q.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (11-NOV-2002) CB&D, E.I. Dupont de Nemours Inc.,			
Experimental Station, Wilmington, DE 19880-0328, USA			
FEATURES			
source			
1. .6334			
/organism="Rhodococcus erythropolis"			
/mol_type="genomic DNA"			
/db_xref="taxon:1833"			
/plasmid="pAN12"			
/note="cryptic plasmid"			
complement(1912..3051)			
/gene="rep"			
complement(1912..3051)			
/gene="rep"			
/note="Rep"			
/codon_start=1			
/transl_table=11			
/product="replication protein"			
/protein_id="AAO31602.1"			
/db_xref="GI:28628257"			
/translation="MTSVSAEHLGSKDRPPVLVSSDKRGIRHRLPKLQLTTSFEN			
ACGRPIGVNGVTIVNGPKSGFGGLSCGKGWICPCGCKVGAHRADEISQVVAHOL			
GTGSVAVMTMRHTAGRLDHLMTGLSAWKAATNGRRMRTEREMGCDGTVRAVEI			
THGKNGHVHVALMFSGDVSENIESFSDAMDFRWTSLKLSIGFAAPLRNSGGLDV			
RKICGEADQVLAAYLTIASGVGMEVSGSGKGRHGNRAPEWIAVAVGDPQAL			
WRPEFGSMGRRAJAWSRGLARAGLGAELTDAQIVOEESAPVAVVAIPARSWMMIR			
TCAPYVGEELGLVEAGATWENLRDLHLHYLRPAADVRPPIISVRK"			
complement(4288..5178)			
/gene="div"			
complement(4288..5178)			
/gene="div"			
/note="Div"			
/codon_start=1			
/transl_table=11			
/product="putative cell division protein"			
/protein_id="AAO31603.1"			
/db_xref="GI:28628258"			
/translation="WDQDTDTPIAGWNELAQPMVMDIAKDAHWLIQKTRSGKSC			
TYNLAQAGSNPAVRVGVDPDTSVLLAPFVHRPAEPNIEGLNDPDKVLRLVQFVKA			
ESDRIECFNDRRIDKISLSPALPLIILVLEEPGLIEGADQDPATNGLKPADRYAP			
RITSLVRQIAQSAKAGIRMLLAQRAEASIVGNGARNFNAVKMLRVDPEPSVKMLH			
PNATPECALVEGEVFGGFFDQPLGRKQMIRTVRVGEYSTVASTYENADLAYEALN			

IDRAQRTIASYPHLDIG"1

ORIGIN

Alignment Scores:

Pred. No.: 5,35e-122 Length: 6334

Score: 2005.00 Matches: 379

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 1 Gaps: 0

US-10-007-527A-2 (1-379) x AV178757 (1-6334)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

DB 3051 ATGACCGGTAAGTCTGACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTCG 2992

QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40

DB 2991 TCCGATAAGCGCGCATCCGGCAGCACTGGACCCCAAACTTCAACAAATCACACGTC 2932

QY 41 GluThrPheAsnAlaCysGlyArgProLysLeuSerGlyValAsnGlyValThrIleValAsn 60

DB 2931 GAAACATTTAACGCTGTGGCGCGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTC 2872

QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80

DB 2871 GGTCGGAAGGTTCTGGATTCGGAGGCTTCGTTCTCGGAAGGGCTGGATTCGCCCC 2812

QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysLeuGlnIleThrThrSer 100

DB 2811 TGCTGTGGGAAAGTCGGTGCACATCGTCAGCAAGAAATTTCTCAAGTTGTGTCAT 2752

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetArgHisThrAlaGlyGln 120

DB 2751 CAACTCGGACTGGATCTGTTGGCATGGTACGATGACCATGACCATGACGTGTCAG 2692

QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

DB 2691 CGGCTCCACGACCTATGACCTGGACTTCGGCAGCTTGGAAAGCTGGACCAACGGTCGT 2632

QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160

DB 2631 CGTTGGCGTACGGAACGTGAATCTACGCTGCGACGATACGTCGCGTGTGTAATC 2572

QY 161 ThrHisGlyLysGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180

DB 2571 ACTCAGGAAACCGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTCACTGGTGAC 2512

QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200

DB 2511 GTGAGTGAGAACATCTCTCGAATCTTCTCGATGCGATGTTGATCGGTGACATCCAAA 2452

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220

DB 2451 CTCGTATCTCTGGGATTTGTCGGCCACTACGTAATTCGGTGGTCTCGATGACGAAAG 2392

QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240

DB 2391 ATCGGCGGTGAAGCTGATCAAGTTCTGCTGCGTATCTGACGAAATTCGATCTGGCGTT 2332

QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260

DB 2331 GGTATGGAGTGTGTAGTGGCAGCGGAAAGTGTGTCGATGCGACCGTGCACCTGG 2272

QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280

DB 2271 GAAATCGCTGTGATGAGTGGCGGGATCCACACGCTTGAACTGTGGCAGAAATTT 2212

QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAla 300

DB 2211 GAGTTGGTTCGATGGGACGTGGGCAATCGCGTGTGTCGATGCGTGGTGGCGGAGCT 2152

QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320

DB 2151 GGTCTTGGGCGAGAACTAACGATCTCAGATCGTTGAGCAGGAAGATCTCCCGGTC 2092

QY 321 MetValAlaIleLeuProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340

DB 2091 ATGTTTGGATCATTTCCGGCGGATCGTGGGATGATTCGACATGTGGCGCTTACGTC 2032

QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360

DB 2031 TTCGGCGAGATCTCGGACTCGTCAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 1972

QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379

DB 1971 TTGCATTATGATTTGCCGCGAGCGATGTGGCGCCCCGATAATATCGTTGCGAAG 1915

RESULT 3

AX548646/c

LOCUS AX548646 6334 bp DNA linear PAT 27-NOV-2002

DEFINITION Sequence 5 from Patent WO20205709.

ACCESSION AX548646

VERSION AX548646.1 GI:25813614

KEYWORDS

SOURCE Rhodococcus erythropolis

ORGANISM Rhodococcus erythropolis

REFERENCE 1

AUTHORS Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.

TITLE Rhodococcus cloning and expression vectors

JOURNAL Patent: WO 0205709-A 5 18-JUL-2002;

E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES

source

1. .6334

Location/Qualifiers

/organism="Rhodococcus erythropolis"

/mol_type="unassigned DNA"

/strain="AN12"

/db_xref="taxon:1833"

ORIGIN

Alignment Scores:

Pred. No.: 5,35e-122 Length: 6334

Score: 2005.00 Matches: 379

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-007-527A-2 (1-379) x AX548646 (1-6334)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

DB 3051 ATGACCGGTAAGTCTGACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTCG 2992

QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40

DB 2991 TCCGATAAGCGCGCATCCGGCAGCACTGGACCACTTCAACAAATCACACGTC 2932

QY 41 GluThrPheAsnAlaCysGlyArgProLysLeuSerGlyValAsnGlyValThrIleValAsn 60

DB 2931 GAAACATTTAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTC 2872

QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80

DB 2871 GGTCGGAAGGTTCTGGATTCGGAGGCTTCGTTCTCGGAAGGGCTGGATTCGCCCC 2812

QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysLeuGlnIleThrThrSer 100

DB 2811 TGCTGTGGGAAAGTCGGTGCACATCGTCAGCAAGAAATTTCTCAAGTTGTGTCAT 2752

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetArgHisThrAlaGlyGln 120

DB 2751 CAACTCGGACTGGATCTGTTGGCATGGTACGATGACCATGACCATGACGTGTCAG 2692

QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

DB 2691 CGGCTCCACGACCTATGACCTGGACTTCGGCAGCTTGGAAAGCTGGACCAACGGTCGT 2632

QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160

DB 2631 CGTTGGCGTACGGAACGTGAATCTACGCTGCGACGATACGTCGCGTGTGTAATC 2572

QY 161 ThrHisGlyLysGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180

DB 2571 ACTCAGGAAACCGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTCACTGGTGAC 2512

QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200

DB 2511 GTGAGTGAGAACATCTCTCGAATCTTCTCGATGCGATGTTGATCGGTGACATCCAAA 2452

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220

DB 2451 CTCGTATCTCTGGGATTTGTCGGCCACTACGTAATTCGGTGGTCTCGATGACGAAAG 2392

QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240

DB 2391 ATCGGCGGTGAAGCTGATCAAGTTCTGCTGCGTATCTGACGAAATTCGATCTGGCGTT 2332

QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260

DB 2331 GGTATGGAGTGTGTAGTGGCAGCGGAAAGTGTGTCGATGCGACCGTGCACCTGG 2272

QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280

DB 2271 GAAATCGCTGTGATGAGTGGCGGGATCCACACGCTTGAACTGTGGCAGAAATTT 2212

QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAla 300

DB 2211 GAGTTGGTTCGATGGGACGTGGGCAATCGCGTGTGTCGATGCGTGGTGGCGGAGCT 2152

Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaThrAsnGlyArg	140
Db	2691	CGGCTCACGACCTATGGATGGACATTCGCGAGCCTGGAAAGCTGCGACCAACGGTCGT	2632
Qy	141	ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160
Db	2631	CGTTGGCGTACGGAACTGAAATGTACGGCTGCGAGCGATGTCGCGCTGTGGAATC	2572
Qy	161	ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	2571	ACTCACGGAAAAACGGCTGGCAGCTCCAGCTTCACGGCTTACTCATGTTCAAGTGGTAC	2512
Qy	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200
Db	2511	GTGAGTGGAGAACATCTCTCGAATCCTCTCGGATCGGATGTTTCGATCGGTGACATCCAAA	2452
Qy	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220
Db	2451	CTCGATCTCTGGGATTGTGTCGCCCACTACGTAATTCGGGTGTTCTCGATGTCGAAAG	2392
Qy	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal	240
Db	2391	ATCGCGGTGAGCTGATCAAGTTCTCGCTCGGTATCTGACGAAAAATTGCATCTGGCGTT	2332
Qy	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260
Db	2331	GGTATGGAGGTGTGTAGTGGCGGACGGAAAAAGTGGTCGACATGGCAACCGTGCAACCTGG	2272
Qy	261	GluIleAlaValAspAlaValGlyCysAspProGlnAlaLeuGluLeuTrpArgGluPhe	280
Db	2271	GAATTCGCTGTGTATGTCAGTGGCGGGATCCACAGGCTTGGAACTGTGGCGAGATTT	2212
Qy	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	2211	GAGTTTGGTTTCATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTGGTGGCGAGATTT	2152
Qy	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal	320
Db	2151	GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTGGAGCGGACTTGGGAAAAATCTTGGCCCGGTC	2092
Qy	321	MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTrpVal	340
Db	2091	ATGGTTGGCATCATTCGGCGCGATCGTGGATGATGATTCGGAATTTGGCGCTTACGTC	2032
Qy	341	PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis	360
Db	2031	TTCGGCGAGATCTCGGACTCGTCGAAGCTGGCGGACTTGGGAAAAATCTTGGTATCAC	1972
Qy	361	LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys	379
Db	1971	TTGCATTATCGATTGCGCGAGCGATGTGCGGCCCCCGGATAATATCGGTTTCGCAAG	1915
RESULT 4			
AX548648/c			
LOCUS			
Sequence 7 from Patent W002055709.			
DEFINITION			
ACCESSION			
VERSION			
AX548648.1 GI:25813616			
KEYWORDS			
SOURCE			
ORGANISM			
artificial sequences; vectors.			
REFERENCE			
1			
Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.			
Rhodococcus cloning and expression vectors			
Patent: WO 02055709-A 7 18-JUL-2002;			
E.I. DU PONT DE NEMOURS AND COMPANY (US)			
FEATURES			
source			
1. .9652			
/organism="Shuttle vector pRHB17"			
/mol_type="unassigned DNA"			
/db_xref="taxon:214938"			
ORIGIN			
Alignment Scores: 8.89e-122 Length: 9652			
Pred. No.: 2005.00 Matches: 379			
Score: 100.00% Conservative: 0			
Percent Similarity: 100.00% Mismatches: 0			
Best Local Similarity: 100.00% Indels: 0			
Query Match: 100.00% Gaps: 0			
DB:			
US-10-007-527A-2 (1-379) x AX548648 (1-9652)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer	20
Db	5052	ATGACCAGCGTAAGTGTGAACACCTTTTCGGCAAGACCGGCTCCGCTCGTCTCGTCG	4993
Qy	21	SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer	40
Db	4992	TCCGATTAAGCGCGCATCCGGCACGAACCTCGACCCCAACTTCACCAAAATCACCACTCA	4933
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	4932	GAACAATTTAACGCTGTGGCCGCCGATTTCTCGCGTGAACGCTGTGACCATTTGTCAAC	4873
Qy	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrIleCysPro	80
Db	4872	GGTCCGAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC	4813
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100
Db	4812	TGCTGTGCGGAAAGTCGGTGCACATCGTGCACGAAATTTCTCAAGTTGTGTGCTCAT	4753
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	4752	CAACTCGGAGCTGGATCTGTTGCCATGTTGTCAGTACCATTCGCCCATACAGCTGTGTCAG	4693
Qy	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140
Db	4692	CGGCTCCACGACCTATGAGCTGGACTTTTCGGCAGCCTGGAAAGCTGCGACCAACCGTCTG	4633
Qy	141	ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160
Db	4632	CGTTGGCGTACGGAACTGTAATGTACGGCTGCGACGATACGTCGCGCTGTGTAATC	4573
Qy	161	ThrHisGlyLysAsnGlyTyrHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	4572	ACTCACGAAAAAACCGCTGGCAGCTCCAGCTTCACGCGTACTCATGTTCACTGCTGAC	4513
Qy	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200
Db	4512	GTGAGTGAAGACATCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGGACTTCCAAA	4453
Qy	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220
Db	4452	CTCGTATCTCTGGGATTGTCTCGGCCACTACGTAATTCGGGTGGTCTCGATGTACGAAG	4393
Qy	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal	240
Db	4392	ATCGCGGTGAAGCTGATCAAGTTCTCGCTCGTATCTGACGAAAAATTCATCTGGCGTT	4333
Qy	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260
Db	4332	GGTATGGAGTTTGGTAGTGGCGACGAAAAAGTGGTGCACATGGCAACCGTGCACCCCTGG	4273
Qy	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280
Db	4272	GAATTCGCTGTGTATGTCAGTGGCGGGATCCCAAGCGTTTGAACCTGTGGCGAGATTT	4213
Qy	281	GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	4212	GAGTTTGGTTTCATGGGACGTCGGGCAATCGCGTGTCTCCGTTGGATTGCTGCCAGCT	4153
Qy	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal	320


```

Db      4152  GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGAGAAATCGCCCGGTC 4093
Qy      321  MetValAlaIlelleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db      4092  ATGGTTGGGATCATTCCTCGGCGCGATCGTGGGATGATGATTCGGACTGTGCGCCTTACGTC 4033
Qy      341  PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db      4032  TTCGGCGAGATCCTCGGACTCGTCGAAGCTGGCGGCACTTGGGAAATCTTCGTGATCAC 3973
Qy      361  LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
Db      3972  TTGCATTATCATGTTGCCGCGAGGATGTGGCGCCCGCGATAATATCGTTGGCAAG 3916

RESULT 5
AX548647/c
LOCUS      AX548647      11241 bp      DNA      linear      PAT 27-NOV-2002
DEFINITION Sequence 6 from Patent WO02055709.
ACCESSION  AX548647
VERSION     AX548647.1  GI:25813615
KEYWORDS   .
SOURCE      Shuttle vector pRHR17
ORGANISM    Shuttle vector pRHR17
            artificial sequences; vectors.
REFERENCE   1
AUTHORS     Bramucci,M.G., Cheng,Q., Kostichka,K.N. and Tomb,J.F.
TITLE       Rhodococcus cloning and expression vectors
JOURNAL     Patent: WO 02055709-A 6 18-JUL-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES    Location/Qualifiers
             1..11241
             /organism="Shuttle vector pRHR17"
             /mol_type="unassigned DNA"
             /db_xref="taxon:214938"

ORIGIN
Alignment Scores:
Pred. No.:      1.07e-121      Length:      11241
Score:          2005.00      Matches:      379
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6              Gaps:        0

US-10-007-527A-2 (1-379) x AX548647 (1-11241)
Qy      1  MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
Db      6641  ATGACCAGCGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTCG 6582
Qy      21  SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
Db      6581  TCCGATAAGCGCGCATCCGGCAGCAACTCGACCCCAAACTTCAACAAATCACACGTC 6522
Qy      41  GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
Db      6521  GAAACATTAAACGCTGTGCGCGCGCATTTCTGGCGTGACCGTGACCATGTGCAAC 6462
Qy      61  GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db      6461  GGTCCGGAAGGTTCTGGATTCGGAGGCTTCGTTCCTCGCGAAGGGCTGATCTGCCCC 6402
Qy      81  CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValValAlaHis 100
Db      6401  TGTGTGCGGAAAGTCGGTGACATCGTGCAGACAAATTTCTCAAGTTGTGTGTCAT 6342
Qy      101  GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db      6341  CAACTCGGACTGGATCTGTGGATGTGGATGACCATGACCATGCGCCATACAGCTGTGTC 6282
Qy      121  ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

```

```

Db      6281  CGGCTCCACGACCTATGACACTTTCGGCAGCGCTCGAAAGCTGCGAACCAACGGTCGT 6222
Qy      141  ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db      6221  CGTTGGCGTACCGAAGCGTAAATGTACGGCTGCGACGATACGTGCGCGCTGTGGAATC 6162
Qy      161  ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db      6161  ACTCAGGAAACACGGCTGCGACGTCCACGCTTTCACGCGCTACTCATCTTCAGTGGTGAC 6102
Qy      181  ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db      6101  GTGAGTGAGACATCTCTCGAATCTTCTCGGATCGGATGTTTCGATCGGTGACATTCCAA 6042
Qy      201  LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db      6041  CTCGTATCTCTCGGATTTGCTCGCCCACTAGTAATTCGGGTGCTCTCGATGTACGAAAG 5982
Qy      221  IleGlyGlyGluAlaAspGlnValLeuAlaIleTyrLeuThrLysIleAlaSerGlyVal 240
Db      5981  ATCGCGGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTCATCTGGCGTT 5922
Qy      241  GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
Db      5921  GGTATGAGGTTGTAGTGGCGCGAAAGTGTGACATGGCAACCGTGCACCCCTGG 5862
Qy      261  GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db      5861  GAAATCGCTGTGATGATGAGTGGCGCGGATCCACAGCGTTGGAACTGTGCGGAGAAATT 5802
Qy      281  GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db      5801  GAGTTGCTGATGAGGACCTCGGGCAATCGCGTGGTCCGTTGATTTGCGTGGCGAGCT 5742
Qy      301  GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluGluSerAlaProVal 320
Db      5741  GGTCTTGGGCGAGNACTTAAAGATGCTCAGATCGTTGAGCAGGAAGAAATTCGCCCGTC 5682
Qy      321  MetValAlaIlelleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db      5681  ATGGTTGCGATCATTCCTCGCGCGCATGCTGGATGATGATTCGACTTGTGCGCTTACGTC 5622
Qy      341  PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db      5621  TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGCGACTTGGGAAATCTTCGTGATCAC 5562
Qy      361  LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
Db      5561  TTGCATTATCGATTGCGCGAGCGGATGTCGGCCCCCGGATAATATCGTTCCGAAG 5505

RESULT 6
AX180162/c
LOCUS      AX180162      11241 bp      DNA      circular      SYN 23-JUL-2003
DEFINITION Shuttle vector pRHR17, complete sequence.
ACCESSION  AX180162
VERSION     AX180162.1  GI:30313707
KEYWORDS   .
SOURCE      Shuttle vector pRHR17
ORGANISM    Shuttle vector pRHR17
            artificial sequences; vectors.
REFERENCE   1 (bases 1 to 11241)
AUTHORS     Kostichka,K., Tao,L., Bramucci,M., Tomb,J.-F., Nagarajan,V. and
            Cheng,Q.
TITLE       A small cryptic plasmid from Rhodococcus erythropolis:
            characterization and utility for gene expression
JOURNAL     Appl. Microbiol. Biotechnol. 62 (1), 61-68 (2003)
MEDLINE     22718480
PUBMED      12835922
REFERENCE   2 (bases 1 to 11241)
AUTHORS     Kostichka,K., Tao,L., Bramucci,M., Tomb,J.-F., Nagarajan,V. and
            Cheng,Q.
TITLE       Direct Submission
JOURNAL     Submitted (14-NOV-2002) CR@D, E. I. Dupont de Nemours Inc.,

```

Experimental Station, Wilmington, DE 19880-0328, USA

FEATURES

Location/Qualifiers

1..11241
 /organism="Shuttle vector pRHBR17"
 /mol_type="genomic DNA"
 /db_xref="taxon:214938"
 /focus
 /note="Rhodococcus - Escherichia coli shuttle vector; Sep I linearize pAN12 cloned into the PvuII site of pBR328"
 join(9925..11241,1..3590)
 /organism="Cloning vector pBR328"
 /mol_type="genomic DNA"
 /db_xref="taxon:221946"
 3591..9924
 /organism="Rhodococcus erythropolis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1833"
 /plasmid="pAN12"
 /note="cryptic plasmid"

ORIGIN

Alignment Scores:
 Pred. No.: 1.07e-121 Length: 11241
 Score: 2005.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-10-007-527a-2 (1-379) x AY180162 (1-11241)

Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
 Db 6641 ATGACCAAGCTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCGCTCGTGTGCG 6582
 Qy 21 SerAspLysArgGlyLysValArgHisGluLeuArgProLysLeuGlnInleThrThrSer 40
 Db 6581 TCCGATAGCCGGCATCCGACGACGACTGGACCCCAACTTCAACAAATCACCACGTCA 6522
 Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
 Db 6521 GAAACATTTAAGCGCTGTGGCGCGGCGGATTTCTGGCGTGAACGGGTGTGACCATTTGTCAC 6462
 Qy 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80
 Db 6461 GGTCCGAAAGGTTCTGGATTCTGGAGCGCTTCTTCTCGGAAAGGGCTGGATCTGCCCC 6402
 Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValValAlaHis 100
 Db 6401 TGCTGTGGGAAAGTCTGGTGACATCTGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 6342
 Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 Db 6341 CAATCTCGGACTGGATCTGTGGATCTGGATCTGGATCTGGATCTGGATCTGGATCTGGAT 6282
 Qy 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTyrLysAlaAlaThrAsnGlyArg 140
 Db 6281 CGGCTCCACGACCTATGACATGAGCTTTCGGAGCGCTTTCGGAGCGCTTTCGGAGCGCT 6222
 Qy 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 Db 6221 CGTTGGCGTACGGAACGTGAATGTACGGCTGCGACGAGTACGTGCGGCTGTGAATC 6162
 Qy 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 Db 6161 ACTCAGCAAAAGCGGTGGCAGCTCCACGCTTCCAGCGCTTCTATGTTTCACTGCTGCTG 6102
 Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
 Db 6101 GTGAGTGAGAACATCTCTGAATCTCTCGGATGATGATGATGATGATGATGATGATGAT 6042
 Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLeuAspValArgLys 220

Db 6041 CTCGATCTCTGGGATTTGCTGGCCACTACGTAATTCGGGTGCTCTCGATACGAAAG 5982
 Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 Db 5981 ATCGCGCGTGAAGCTGATCAAGTTCTCGCTGCGCTATCTGACGAAATTTGATCTGGCGTT 5922
 Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
 Db 5921 GGTATGAGGTTGGTAGTGGCGACGCGGAAAGTGGTGCACATGCGCAACCTGCACCTCG 5862
 Qy 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
 Db 5861 GAAATCGCTGTTGATGCGTGGCGGGATCCACAGCTTGGAACTGTGGCGAGAAATTT 5802
 Qy 281 GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
 Db 5801 GAGTTTGGTTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGGATTCGCGCCGAGCT 5742
 Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 Db 5741 GGTCTTGGGCGAGACTTACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC 5682
 Qy 321 MetValAlaIleleProAlaArgSerTyrMetMetIleArgThrCysAlaProTyrVal 340
 Db 5681 ATGTTTCCGATCATTCGCGCGGATCGTGGATGATGATGATGATGATGATGATGATGATG 5622
 Qy 341 PheGlyGluLeuLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
 Db 5621 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562
 Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProIleLeuSerValArgLys 379
 Db 5561 TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCCCGGATATATCGTTCCGCAAG 5505
 RESULT 7
 APU83788 2439 bp DNA circular BCT 15-JUN-1998
 LOCUS Actinomyces pyogenes plasmid pAP1, complete plasmid sequence.
 DEFINITION U83788
 ACCESSION U83788
 VERSION U83788.1 GI:1805288
 KEYWORDS Arcanobacterium pyogenes
 SOURCE Arcanobacterium pyogenes
 ORGANISM Arcanobacterium pyogenes
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Actinomycineae; Actinomycetaceae; Arcanobacterium.
 REFERENCE 1 (bases 1 to 2439)
 AUTHORS Billington, S.J., Jost, B.H. and Songer, J.G.
 TITLE The Arcanobacterium (Actinomycetes) pyogenes plasmid pAP1 is a member of the pJ101/pJVI family of rolling circle replication plasmids
 JOURNAL J. Bacteriol. 180 (12), 3233-3236 (1998)
 MEDLINE 98292760
 PUBMED 9620977
 REFERENCE 2 (bases 1 to 2439)
 AUTHORS Billington, S.J., Jost, B.H. and Songer, J.G.
 TITLE Direct Submission
 JOURNAL Submitted (06-JAN-1997) Veterinary Science, University of Arizona, 1117 East Lowell Street, Tucson, AZ 85721, USA
 FEATURES
 source
 1..2439
 /organism="Arcanobacterium pyogenes"
 /mol_type="genomic DNA"
 /strain="BBR1"
 /db_xref="taxon:1661"
 /plasmid="pAP1"
 rep_origin 1..110
 /note="single stranded origin; putative"
 rep_origin 39..44
 /note="site of single strand initiation; putative"
 gene 221..1600
 /gene="rep"
 CDS 221..1600
 /gene="rep"
 /function="replication protein"

```

/codon_start=1
/transl_table=11
/product="Rep"
/protein_id="AAC46399.1"
/db_xref="GI:1805289"
/translation="MNRLSRTALSLPARQIQKIPVPAAGGRSLKSPFGMTATWSARGG
ASDERRDRSRQIPSNRRGRSATHPLGNTVLTFVPSNESKKTAKSRSEVERELRDG
LAETSTESVRCGRVPVAPLVSRLAKSDKGAGYGLHTCGSVWACPVCSAKIAARR
KTLQOVVHVAVKHGVTVMILTQRHHKQGLKHLWDALSTANWVTSGRRWIEFKR
QFGLGVVVRANEITHGHGWHVSHVLIISKKDPLTSTFVYQKQRRRLFPYPEIYM
SSDFIARWGAIAKHGVDLFRDGGDLDTVAKDARAIGNVYSKMQTSTDAISSEVTL
GGFKABGNRTTPOILADILSLGDVLDLWKVEYKASFGRRALTWSKGLRDWANUG
VEGSDIEIASIEEIGDEAIALFTHDARQVRRFGAELLDVTESGRRAAAYRWLDFRI
DWGLPPKIE"
1673. 1676
RBS
CDS
1685. .2023
/note="ORF112"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC46400.1"
/db_xref="GI:1805290"
/translation="MFFIGYRTKFOPIAPCKSPFIPDGNKRTGFSSEKMRTOAIT
EDGLVELTSIPQFVQLIDDAVKSRLELFEKVNLAVLPRNGGISTYLTGEAVHV
PAFVVVEVS"
2010. .2013
RBS
CDS
2020. .2409
/note="ORF129"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC46401.1"
/db_xref="GI:1805291"
/translation="MMSEPYGISSNAELQLVFRPAKREARLYMSRLKKELESKHS
TPTEALESCEKLYWSEFPEPLVDVILHEVGECCACGTNSAVARSLSRAPFARFVQS
LENGGRNPAQLQDQVAGMRKTIGTR"
2378. .2406
rep_origin
/note="double stranded origin; putative"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 7.38e-31 Length: 2439
Score: 609.50 Matches: 148
Percent Similarity: 51.47% Conservative: 62
Best Local Similarity: 36.27% Mismatches: 133
Query Match: 30.40% Indels: 65
DB: 1 Gaps: 15

```

US-10-007-527a-2 (1-379) x APU83788 (1-2439)

```

QY 5 SerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSerSerAsp----- 22
Db 419 TCAGGACCATCCCTTGGCAATACGGTATTAACTTTTCAGTATCAACAGAGCAAG 478
QY 23 -----LysArgGlyLleArgHisGluLeuArgProLysLeuGlnGlnle 37
Db 479 AAAACAGCTAAATCTCGCGTCTTGAGAGATACAACTCAGACCGATTAGCCGAAATC 538
QY 38 ThrThrSerGluThrPheAsnAlaCysGlyArg---ProIleSerGlyValAsnGlyVal 56
Db 539 TCGACCATTTGAGTCGCTCGGGAAGTGTGGCGGTGCGCGGTGCGCACCTCTCGTCTGTTG 598
QY 57 ThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGly 76
Db 599 CGAGCAAAATCTGACGGTAAGCGCGGATATGGTGTTCGACACTTGTGGAAGCGTC 658
QY 77 TrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGln 96
Db 659 TGGCGTGCCTCAGTCTGTAGCGGAATAATCGCGCTCGCGCAAAACCGACCTCAACAG 718
QY 97 ValValAlaHisGlnLeuGlyThrGly---SerValAlaMetValThrMetThrMetArg 115
Db 719 GTCGTTGACACCGCCGTAAACACGGAATGACCGTCTCAATGCTTACGCTACCCAGCGT 778

```

```

QY 116 HisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTrpLysAla 135
Db 779 CACCACAAGGACAAAGGCTAAACACCTCTGGGACGCTTGTGACGCGCATGGAATCGC 838
QY 136 AlaThrAsnGlyArgArgTTPArgThrGluArgGluMetTyrGlyCysAspGlyTyrVal 155
Db 839 GTTACCTCTGTGCTGCTGTGATTGAGTTCAAGAGCAAAATTGTGTTTAGTCGGTTATGTT 898
QY 156 ArgAlaValGluIleThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeu 175
Db 899 CGAGCCCAATGAATTACTCATGGAACACGCGTGGCATGTGCATTCCTCATGTTCTGATT 958
QY 176 MetPheSerGlyAspValSerGluAsnIleLeuGluSerPhe----- 189
Db 959 ATTTCCGAGAAAGAC-----CCGCTGACCAAGCACGTTTGTCTATCAACGCAACAA 1009
QY 190 -----SerAspAlaMetPheAsp 195
Db 1010 GAGCGCGCGCGCTTCCCTACCCCGCCAGAGATTATATGTCATCCGATTTTCATTGCTGAA 1069
QY 196 ArgTTPThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGly 215
Db 1070 CGGTGGGAAGCTCGCCTTGGGAAGCACGCGGTGATT---CTCCCGAATTCGGAGGC 1126
QY 216 Leu-----AspValArgLysIleGlyGlyGluAlaAspGlnValLeu 229
Db 1127 TTGGACTGGACCGCTTCGAAAGACGCGCGAGCCATCGGCAAC----- 1168
QY 230 AlaAlaTyrIleuThrLysIle-----AlaSerGlyValGlyMetGluValGlySer 246
Db 1169 ---TATGTCAGCAAGATGTCAGACGTCACACAGCGGATTAGTCGGAAGTCACGTTG 1222
QY 247 GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAla 266
Db 1223 GCGCGCTTCAAAAAGCCCGAAACGGAACAGACGCGCTTCAGATATCTCGCGGATATC 1282
QY 267 ValGly---GlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPheGlySerMet 285
Db 1283 CTTTCGCTCGCGCATGTCGACGACCTCAAGCTCTGGAAGAATAATGAGAAAGCTTCGTT 1342
QY 286 GlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGlu 305
Db 1343 GGACGCGCTGCACTTACATGCTCGAAAGGCGCTCAGAGATTGGGCAAAATCTCGCGTTGAA 1402
QY 306 LeuThrAspAlaGlnIleValGluGlnGluSerAlaProValMetValAlaIleIle 325
Db 1403 CAGTCGCGACGAAGAGATTGCC---TCTGAGAAATCGGGGACGACCAATAGCGCTATT 1459
QY 326 ProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrValPheGly-----Glu 343
Db 1460 AGCATGACGCTTGGCGTCAGGTGCGACGT-----TTTGGAGCGCGCTGAA 1504
QY 344 IleLeuGlyLeuValGluAlaGlyAla-----ThrTrpGluAsnLeuArg 358
Db 1505 CTACTCGATGTCACGAATCCGAGGTCTGCGCGCGCTTACCGCTGTTGGATTTTAGG 1564
QY 359 AspHisLeuHisTyrArgLeuPro 366
Db 1565 GAA---ATTGATTGTCATTGCGCT 1585

```

RESULT 8

```

AY150274 3540 bp DNA circular BCT 02-JAN-2003
LOCUS Propionibacterium granulosum cryptic plasmid pPG01, complete
DEFINITION sequence.
ACCESSION AY150274
VERSION AY150274.1 GI:27465054
KEYWORDS Propionibacterium granulosum
SOURCE Propionibacterium granulosum
ORGANISM Propionibacterium granulosum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Propionibacterineae; Propionibacteriaceae; Propionibacterium.

```

REFERENCE	1 (bases 1 to 3540)	4 ValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSerSerAspLys 23
AUTHORS	Farrar, M.D. and Holland, K.T.	2094 GTACACGGCAGCAACTTTTCGGTATCAGATTCCCTCTGTAGA-AAGCGATGAGGCG 2152
TITLE	Isolation and characterisation of a cryptic plasmid from the human skin commensal <i>Propionibacterium granulosum</i>	
JOURNAL	Unpublished	24 ArgGlyIleArgHisGluLeuArgProLysLeuGln-----GlnIleThrThrSerGlu 41
REFERENCE	2 (bases 1 to 3540)	2153 CGTCGCCATCGCTATCAGATGAGGATGGTCTACGAAATCCGCGAGGTCATCGCGCTGAG 2212
AUTHORS	Farrar, M.D. and Holland, K.T.	
TITLE	Direct Submission	42 ThrPheAsnAlaCysGlyArg---ProIleSerGlyValAsnGlyValThrIleValAsn 60
JOURNAL	Submitted (13-SEP-2002) Skin Research Centre, Division of Microbiology, University of Leeds, Leeds LS2 9JT, UK	2213 CGCGTTCCGAAAGTCGCGGCGCAGTCGCGTTTCG-----CAACGGATCCGCTTGATGGCG 2266
FEATURES	Location/Qualifiers	
source	1..3540	61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80
	/organism="Propionibacterium granulosum"	2267 GGTTCATGGTGGTCCGCTTATGCGGGTTTGGCGACGTGCGGAAGTGTGGCTTGCCT 2326
	/mol_type="genomic DNA"	
	/strain="PF283"	81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValValAlaHis 100
	/db_xref="taxon:33011"	2327 GTCTGTGCGGCAAAAGATTTCGCGCACCGTCGTGATGAGCTGGCGCGCTGTTCAGGTT 2386
CDS	/plasmid="cryptic plasmid pPG01"	101 GlnLeuGlyThrGly---SerValAlaMetValThrMetThrMetArgHisThrAlaGly 119
	74..280	2387 GCGGTTGGACTCGGCTTCAAGGTGTGATGCTGACGCTTACTCAACGTCACTGCTGGT 2446
	/codon_start=1	120 GlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGly 139
	/transl_table=11	2447 CAGGATCTCCCGCAGCTGTGGCGTCTGCTCAGTGGGTGGATGTGTGCACGAGTGGT 2506
	/product="Orf3"	140 ArgArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGlu 159
	/protein_id="AA078121.1"	2507 CGACGGTGGCAGGAATTTTCGCTCAGCTCGCGTCCAGGGATGGGTCAAGGCAGATTGAA 2566
	/db_xref="GI:27465055"	160 IleThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGly 179
gene	/translation="MGQYGLNRGFYNGSHNELIGYPMNGRVIWVWDPALLSAMPLSSGFSFGHVVEPRAEVFHRCADHD"	2567 GTCACCATGGGTGCATGGGTGGCAGCTTCACTGCGTGCAGTGCCTCATCTCTAAGCAG 2626
CDS	449..1792	180 Asp-----ValSerGluAsnIle----- 185
	/gene="tra"	2627 GATCCGACTAGCTGTGACATAAGATTGGCATCGCCGCAACAAAGGTTCGCGCGCGGACC 2686
	/note="plasmid transfer protein"	186 -----LeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
	/transl_table=11	2687 CCGTATCCAGAGAGGTACAGAGCCCGAAGACTTTCATCGTCAACGCTGCTCGCGAGGT 2746
	/product="Tra"	201 LeuValSerLeuGly-----PheAlaAlaProLeuArgAsnSerGlyGlyLeuAspVal 218
	/protein_id="AAN78122.1"	2747 TTGAGGAAGCGCGCGCTCGACTTCATCGCC-----GGTAGTGGTGGCTCGATTGG 2797
	/db_xref="GI:27465056"	219 ArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSer 238
	/translation="MAGERTVLTGVYWMRLILPLVLGVAFACWCFSVHLVVIGFAVAALVAFPPVQVPMQVPAVYVYRWWSLAKAFKIESYGGDRIVYRPLQWI RSDRVHLVLRVAPGLASAAYLEKGAEEIQRQRGKSKWRTCIIVRPAEHLDIIFR DQAGBELPAGVSTSNVYVGVKPDGSEVVDLSHPSHILVSGTKSGSKSFVYGLL DOMEHLFVTVAGVPTGILFENELGDGWDGALRSKRITNDADAANVQVLSMITDEND RRIYLLNCHRRDQKGVNDFSDPGRRLLIIVILEYPLGLIERLQNFDSARGARSDDRA SKAGLVGRVAYEAKGVGVLLVLTQRPDAKIIGGPLRAQLTTRVTFQASDGLRMSH PELSEQVQKNSWASGVGFIEADGVIPLTRPRSYRAELTDLHRFASVGVQIDLIQ"	239 -----GlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGly 255
gene	1979..3280	2858 TCCGTCGATGGCTTAGCGAACAGAGCCAGCTGGTGGCGGGTTCAGAGAGGCTCGTAGAGGT 2917
CDS	/gene="rep"	256 AsnArgAlaProTyrGluIleAlaValAspAlaValGly---GlyAspProGlnAlaLeu 274
	/gene="rep"	2918 AATCGGACGCGCTTCCAGATCCTCGAAGATTCCTCGATACCGGCTCGGAGACTGACCTG 2977
	/note="plasmid replication protein"	275 GluLeuTyrArgGluPheGluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArg 294
	/codon_start=1	2978 AGACTCTGGCGTACCTATGTTTCTGCAAGTCATGCGCGTCAAGGCATTCAGCTGGTCCA 3037
	/transl_table=11	295 GlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGln 314
	/product="Rep"	3038 GGTTCGCTGATCGGCTGGCATGGAATCTGAGATGAGCAGTACGACAGCTGCCGCCCA 3097
	/protein_id="AAN78123.1"	315 GluGluSerAlaProValMetValAlaIlelleProAlaArgSerTyrMetMetIleArg 334
	/db_xref="GI:27465057"	3098 GACCAAGTCGGGGAGAGCG---GTCCGCTTCTTGTACCATGACCGCTGGCGGAGATCCGC 3154
	/translation="PWLNIIVTLISKAQPHDLCSHEPCAAERSERQAAAGDRIGTTRATFPVSDSPCKRKNRGRHRYEMRDGLRNPQVMPLEVRKCGAPVPSQRIALMAGHQDLA YGLATCGSWACPVCAAKISAHRRDELARVQVAVGLGVKVSMLTLTQRHHAQDLA FLASLOGSWAVTSGRRWQEFCAQLGVQGVWKAVEVTHGSHGHVHVHVLVLSKQDP TSVPTKIRHRKQRRRTTYPPEVQRPEDFIAERWSRGLRKRGVDFIAGSGGLDWQTA DSGDEALGRVYAKMNSVDGLANEATLGGFKARRGNRTPFQILEDFLDGSETDLR LMRTVVSASHGRKALTWSKGLRDWAGMESEMDQVAAODCQGEAVALFDHDAWRQIR TAGAFLLELELHSGSEGVYAWLKKRLIHYEILPLVPWSTST"	2979 ----- 3037
ORIGIN		
Alignment Scores:		
Pred. No.:	5,28e-28	Length: 3540
Score:	569.00	Matches: 141
Percent Similarity:	51.78%	Conservative: 63
Best Local Similarity:	35.73%	Mismatches: 152
Query Match:	28.38%	Indels: 39
DB:	1	Gaps: 14
US-10-007-527A-2 (1-379) x AV150274 (1-3540)		

```

Db      3155 ACTGCGGTGCGCTTTCTCTCTGACGAGCTGGAGCTCCACGATCCGAGGGCGGTCTAC 3214
QY      353 ThrTrpGluAsnLeuArgAspHisLeuHisTyrArgLeuPro 366
Db      3215 GCGTGGCTGAAGACGCGA--AGAAATCCATTATGAGATACCT 3253

RESULT 9
LOCUS   PFR6662                2051 bp    DNA    circular BCT 08-JUN-1998
DEFINITION Propionibacterium freudenreichii plasmid pLME108 rep gene.
ACCESSION AJ006662
VERSION   AJ006662.1 GI:3212127
KEYWORDS putative; rep gene
SOURCE    Propionibacterium freudenreichii
ORGANISM Propionibacterium freudenreichii
          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
          Propionibacterineae; Propionibacteriaceae; Propionibacterium.
REFERENCE 1
AUTHORS   Dasen,G.H., Miescher,S., Teuber,M. and Meile,L.
TITLE     Molecular analysis of plasmid pLME108, a plasmid isolated from
          Propionibacterium freudenreichii
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 2051)
AUTHORS   Dasen,G.H.
TITLE     Direct Submission
JOURNAL   Submitted (03-JUN-1998) Dasen G.H., Institute of Food Science,
          Laboratory of Food Microbiology, ETH Zurich, LFO G24.2,
          Schmelzbergstr. 9, CH-8092 Zurich, SWITZERLAND
FEATURES             Location/Qualifiers
     source            1..2051
                     /organism="Propionibacterium freudenreichii"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1744"
                     /lab_host="DF2"
                     /plasmid="pLME108"
                     complement(223..1434)
                     /gene="rep"
                     complement(223..1434)
                     /gene="rep"
                     /codon_start=1
                     /transl_table=11
                     /product="putative rep protein"
                     /protein_id="CAA07175.1"
                     /db_xref="GI:3212128"
                     /db_xref="GOA:O69764"
                     /db_xref="TrEMBL:O69764"
                     /translation="MVVDDARARSGVRGCGGSPAVAPLGGALDNHANNVSPSPVAPAA
                     GEVADAKRHRTSVKMRVGLWELNATSLKRVACFGHVAASAVASVGRVCSGDRAGPAGLQSC
                     GSVWACPCVNCNIAVRMGRLEGNATVETVHKGRVAFMTFTVHRSKOSLTATVMDGVA
                     SGWRVTSCKGWTSDLRHGCVGFVRVVEVTHGRNGMVHVLHVLFVLDGFDGALALH
                     RSMFGKRWKGLAAGLGTPLAAQDVQMSATGLDLHARLYLSKAQFGQKIGHETLNS
                     QSKTASALSTWELVDANGLAKEVLMFWFKSGRGRROIQGWAGLRLDGLM
                     VRESDVIAAEKVGVSADTVALITGDWRRLVGQOKLYECLRSCELGCGAGLARWLQK
                     HGTEHLELVEVD"

ORIGIN
Alignment Scores:
Pred. No.:      1..46e-23      Length:      2051
Score:          497.00         Matches:     127
Percent Similarity: 50.80%      Conservative: 63
Best Local Similarity: 33.96%   Mismatches:  144
Query Match:    24.79%         Indels:      40
DB:             1              Gaps:        15

US-10-007-527A-2 (1-379) x PFR6662 (1-2051)

QY      19 ValSerSerAspLysArgGlyIleIleArgHisGluLeuArgProLysLeuGlnIleThr 38
Db      1296 GTGGCTGCTGACAGCGCAACACCGGTTTTCGTCGGTATTGTTGGTGGCATACG 1237
QY      39 ThrSerGluThrPheAsnAlaCysGlyArg---ProIleSerGlyValAsnGlyValThr 57
          :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

```

Db	1236	TCGTTGAAGCGGGTTGGCTTCTCGCGTCTGGTTCGCCGCTCGCGGGTTCGCTTCGGTTGGT	1177
Qy	58	IleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyr	77
Db	1176	GTGGCGGTGCACGACGCGCGCGGGTTCGCTGGTTCAGAGCTCGCATCGGTCTGG	1117
Qy	78	IleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal	97
Db	1116	GCCTGCCCGGTGTGTAAAGCGAAGATCATGGCTCGCGTGGCCCTTGAACTCGGCGGGC	1057
Qy	98	Val-----AlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThr	113
Db	1056	CTCGAGACCTGCAATAAGCAC-----GGTGTGCTGTGGCGTTTCATGACGTTTCACG	1006
Qy	114	MetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaLaTyr	133
Db	1005	GTTTCGGACATCTCGTAAGGATTCCTTGACGCGCTGTTTGGATGTTGCTTCGGGGTGG	946
Qy	134	LysAlaAlaThrAsnGlyArgGlyArgThrGluArgGluMetTyrGlyCysAspGly	153
Db	945	CGGCGAGTCACTTCGGGCACAGGCTGACCTTGACCGCTCCGTACCGGGTCGAGGGC	886
Qy	154	TyrValArgAlaValGluIleThrHisGlyLysAsnGlyTyrHisValHisAla	173
Db	885	TTCTGCGTGTGGTCGAGGTACGCACGGCGCAATGGTGGCATGTGCATCTGCATGTG	826
Qy	174	LeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMet	193
Db	825	CTCGTGTCTGTCGGCGACTTCGGCGAGCTCTG-----CGCTGCACCGGTCATG	772
Qy	194	PheAspArgTyrThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSer	213
Db	771	TTTCGGCGTTGGAAACGCGCGTGTCTCGTGTGGCTCGGACTCCCTTCTCGCGCG	712
Qy	214	GlyGlyLeuAspValArgLysIleGlyGlyGluAlaAsp---GlnValLeuAlaLaTyr	232
Db	711	-----CAAGACGTTTCAGCAGATGAGCGCAGCAACTGGCGCTGGATCACCTCGCGCGCTAC	658
Qy	233	LeuThrLys-----IleAlaSerGlyValGlyMetGluValGlySerGlyAspGlyLys	250
Db	657	CTGAGCAAGCGCGAGTTCGAGGGGAAGATCGGCCACGAGTTGACCAACTCCACAGACAAG	598
Qy	251	SerGlyArgHis-----GlyAsnArgAlaProTyrGluIleAlaValAspAlaValGly	268
Db	597	ACGCGCGGTTCTCGCTCTCGACTCGATCGACTTGGAGGCTCTGGGCGACCGCGCCAC	538
Qy	269	GlyAspProGlnAlaLeuGluLeuTyrArgGluPheGluPheGlySerMetGlyArgArg	288
Db	537	GGCTTCGCGAAGAGGTGGTCTCTGTTTCGAGTGGAGAGGGTAGCCGGGACGCAGA	478
Qy	289	AlaIleAlaTyrSerArgGlyLeuArgAlaArgAlaGlyLeuGlyValaGluLeuThrAsp	308
Db	477	CAGATTGGCTGTCTCGCGGGTTGAGGACCGCTCTCGGCCTCATGGTTGAGGAATCTGAC	418
Qy	309	AlaGlnIleValGluGlnGlu---GluSerAlaProValMetValAlaIleIleProAla	327
Db	417	GACGTGATCGCGCTGAGGAAGTCGGGCTGTGTGTGACATGTCTCGCTTTCATCACGGGT	358
Qy	328	ArgSerTyrMet-----MetIleArgThrCysAla	337
Db	357	GACGATGGCGTCGCTGGTGGGCGACGAGACTTACGAGTGTCTGCGAAGTTGC---301	
Qy	338	ProTyrValPheGlyGluIleLeuGlyLeuValGluAlaGly---AlaThrTyrGluAsn	356
Db	300	-----GAATTGGCGGT-----CAAGCTGGCTTGGCGCGGTGG-----268	
Qy	357	LeuArgAspHis---LeuHisTyrArgLeuProAlaAlaAsp	369
Db	267	CTGCAAGACGACGGAATCGAACACGAAATTGGTGGAGGTGGAC	226
RESULT 10			
SLMPG33A			
LOCUS	SLMPG33A	2661 bp	DNA linear BCT 09-MAY-1994

S.lavendulae miniplasmid pSLG33 sequence.

DEFINITION X69872
 ACCESSION X69872.1 GI:47179
 VERSION miniplasmid.
 KEYWORDS Streptomyces lavendulae
 SOURCE Streptomyces lavendulae
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE 1
 Pelsberg, J., Petricek, M. and Tichy, P.
 Nucleotide sequence of the mini-plasmid pSLG33 from Streptomyces lavendulae-grassierus RIA746
 JOURNAL Nucleic Acids Res. 21 (15), 3582 (1993)
 MEDLINE 93348001
 PUBMED 8346038
 REMARK (sites)

REFERENCE 2 (bases 1 to 2661)
 Pelsberg, J.
 Direct Submission
 AUTHORS Submitted (28-DEC-1992) J. Pelsberg, Inst. of Microbiology, Cell and Molecular Microbiology Div., Videnka 1083, Prague 4, CZECHOSLOVAKIA
 JOURNAL

FEATURES
 source
 1..2661
 /organism="Streptomyces lavendulae"
 /mol_type="genomic DNA"
 /strain="RIA746"
 /db_xref="taxon:1914"
 /dev_stage="mycelia"
 /plasmid="miniplasmid pSLG33"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,55e-20 Length: 2661
 Score: 453.00 Matches: 129
 Percent Similarity: 48.33% Conservative: 59
 Best Local Similarity: 33.16% Mismatches: 160
 Query Match: 22.59% Indels: 41
 DB: 1 Gaps: 16

US-10-007-527A-2 (1-379) x SLMPG33A (1-2661)

Qy 3 SerValSerAlaGluHisLeuSerGlyLysAspArgPro-----ProValLeu 18
 Db 586 AATGTTTCCACCTCCCTCTCTTCCACTCAATCTCCAGCTCAGAGCGGATTCGCC 645
 Qy 19 ValSerSerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnLeuThr 38
 Db 646 GTTCGTAGGCGCGCGGAACGGCGCTTCGCCAGCGGCGAGTCTCTGGACGATCACC 705
 Qy 39 ThrSerGluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGly----- 55
 Db 706 AGCTTTCCTCCGCGTCCGCAAGTCGCTCGGACGACCAATGGAGGCGTCCGTGACGCGCTG 765
 Qy 56 ---ValThrIleValAsnGlyProLysGlySerGlyPheGlyLeuArgSerCysGly 74
 Db 766 GCGGCGATCGTGTCAACGGC---TCCGTCGCTACTGCTCGCGCTTCTGCACTCGCGT 822
 Qy 75 LysGlyTyrIleCysProCysCysAlaGlyLysValGlyValHisArgAlaAspGluIle 94
 Db 823 TCCATCTGGGCTGCGCGTGTCTGTCGGCCAGATCAGGCGCCACCGCGCGACGAGATC 882
 Qy 95 SerGlnValValAlaHisGlnLeuGlyThrGlySerValAla---MetValThrMetThr 113
 Db 883 GCACGAGCGCGCGCCCAAGCACCTCGCGCGCGCGCGCGCGCTGGATGTCTACGTGACG 942
 Qy 114 MetArgGlyThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTyr 133
 Db 943 GCCCGACACAAGAACACACAGAGTTGAGCCCTCTTCGACCGCGTCCGACGCGCTGG 1002
 Qy 134 LysAlaAlaThrAsnGlyArgArgTrpArg-----ThrGlu 145
 Db 1003 CGCAAGCTCTGTCCGCGGAGGCTTGGCGCGGAGACCCCAAGCGGCGAGTCTCTCGCGGAA 1062

Qy 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGly---Lys 164
 Db 1063 CGGACCGGCTCGAGTCCGGGCAACATCCGCTCACTCGAAGTCACTTACGGAACCTCGT 1122
 Qy 165 AsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsn 184
 Db 1123 AACGGTTGGACCCGACCTTCACGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1182
 Qy 185 IleLeuGluSerPheSerAspAlaMetPheAspArg---TrpThrSerLysLeuValSer 203
 Db 1183 CTC-----GCCATCCGCAATGACCGATGGGCAAAACCTGGAGGCGCTCGCTCAAGAAA 1236
 Qy 204 LeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGly 223
 Db 1237 GCCGATTCCGACCGAGCAAGAACAC-----GGCATCACCTGGTCCAAAGTCTC 1284
 Qy 224 GluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGly---ValGlyMet 242
 Db 1285 ACGACCCCAAGAGAGGCGGCGAATACATCCCAAGGCTCAAGAGGAGGAGTCCGAAAC 1344
 Qy 243 GluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGlu--- 261
 Db 1345 GAAATCCGCGGAGCGGACATGAAGCGCGGCTCGGAACTCTGGCCCGCTTCCGAAATG 1404
 Qy 262 IleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGlu 281
 Db 1405 CTCGAATCTCCGCGAGACCGGTGACATGGCAGTGTCCCGATTTGGCAGGATGACGAG 1464
 Qy 282 PheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArg---Ala 300
 Db 1465 AAGGGCACGTTCCCGCTCGCGCGATCACCTGGAGCGGTGCGCTTCGTCGCGAACTCCTC 1524
 Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 Db 1525 GGACGCAAGAGAGAGTCCCGCAAGAACTTCAGCTCGAGGAGATCGGAGGCGGAGACG 1584
 Qy 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
 Db 1585 TGG---GCATCTCTCGCGCGGCGAATCGCTCCGAGCGATCCGCGT---GTCCCGGTC 1638
 Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrp----- 354
 Db 1639 CAGTCCCGCATCTCCGACCGCGAGAAACGCGCGCTTCGTGGAACCTGTCACCTGCTGA 1698
 Qy 355 -----GluAsnLeuArgAsp 359
 Db 1699 CTGCCTACCATCTTCTCTACGAGAT 1723

RESULT 11
 SNA243257 9367 bp DNA circular BCT 02-DEC-1999
 LOCUS Streptomyces natalensis plasmid pSNA1 including spda, kora, spdb2
 DEFINITION and rep genes.
 ACCESSION AJ243257
 VERSION AJ243257.1 GI:6523476
 KEYWORDS KORA protein; rep gene; replication protein; SpdA protein; SPDB2 protein; transfer protein.
 SOURCE Streptomyces natalensis
 ORGANISM Streptomyces natalensis
 Bacteria; Actinobacteria; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE 1
 Mendes, M.V., Aparicio, J.F. and Martin, J.F.
 Complete nucleotide sequence and characterization of pSNA1 from the Pimaricin-producing Streptomyces natalensis that replicates by a rolling circle mechanism
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 9367)
 TITLE Mendes M.V.
 JOURNAL Direct Submission
 Submitted (23-JUN-1999) Mendes M.V., Microbiology, Inbiotec, AV. REAL 1, Leon, 24006, SPAIN

FEATURES
source

```

Location/Qualifiers
1..9367
/organism="Streptomyces natalensis"
/mol_type="genomic DNA"
/db_xref="taxon:68242"
/plasmid="psN1"
/country="South Africa"
266..814
/gene="spda"
266..814
/gene="spda"
/function="putative intramycelial transfer protein"
/notes="ORF4"
/codon_start=1
/transl_table=11
/product="SPDA protein"
/protein_id="CAB62258.1"
/db_xref="GI:6523477"
/db_xref="GOA:Q9RIQ3"
/db_xref="TrEMBL:Q9RIQ3"
/translation="MYVTAPHTGWPSPSAGTVTGMTOGETPRDAGRRICEHCLVKP
VPSLGNKPSVCSNRCQRAYEARQGRFVKQTVNALLRRMMRANATQATSRDNP
EAKSRDAGQATSRDVASAPATRPAPAPRPPVQPPQTAALFEEEDQADEEAP
RPPARRMTDLXQMPKUPGGFG"
complement (1203..1940)
/gene="kora"
complement (1203..1940)
/gene="kora"
/function="putative transcriptional regulator"
/notes="ORF1"
/codon_start=1
/transl_table=11
/product="KORA protein"
/protein_id="CAB62259.1"
/db_xref="GI:6523478"
/db_xref="GOA:Q9RIQ2"
/db_xref="TrEMBL:Q9RIQ2"
/translation="MPKIERASPPYQMOWTHFRTRIMSGELSPGDALPTDRGLAEFWG
VSRATQALTVLTKTEGLIEAVQSGTRVQSAPLHNGHDAVSRVTRTGRVYAGBY
ARITSALEPAPADVAALGIEGAPAIRIRVYNASDVALSMTSLWDGALSSTAP
ALVSTRIRGTRWYLETKTGVVQGDVVRVRLATDEBAQLLSLEQPAAVKLTRTI
LRDEGTLLEYGVSVTAGRESIYEYTV"
2096..3025
/gene="spdb2"
2096..3025
/gene="spdb2"
/function="putative intramycelial transfer protein"
/notes="ORF2"
/codon_start=1
/transl_table=11
/product="SPDB2 protein"
/protein_id="CAB62260.1"
/db_xref="GI:6523479"
/db_xref="GOA:Q9RIQ1"
/db_xref="TrEMBL:Q9RIQ1"
/translation="MFVRVPHRPAGADPAQLERILRLGLTVLVVVAVGCLFATATNV
TFALDHGVSRTAWLLDPLVASALGTVLVDGLLSEFVRPGMASLLRWFLATW
LNMWCSLWPTGPGIPHLPDAGLLHSVLPLVLAALITTRRIRVSKIAVID
EADTCWPTVDTPMPASAPVRSEPLTASAPAAPAPVRPQRELEVICGABQVFR
PALPAPVPSPKPEPAPVPFNGADDPHVAIVRAVWMDGELSTROTAEVRVGLSKTTVA
NWRLEAAHSGQRPAFQPLPGAGQTVLQGVAA"
6656..8086
/gene="rep"
6656..8086
/gene="rep"
/notes="ORF3"
/codon_start=1
/transl_table=11
/product="replication protein"
/protein_id="CAB62261.1"
/db_xref="GI:6523480"
/db_xref="GOA:Q9RIQ0"
/db_xref="TrEMBL:Q9RIQ0"

```

```

/translation="MTTDKALSRCRYALGGGVTPKVTGGTAYLAGLATCGKVHICP
CCGAKIRSARTVELQAGTAWEDIGINGPKIRPARAAWKDINDRTAEDWAAIEED
GLOAGGLGMLTLTMRHYSRHTLAEILVTQODAWKALQONAGRDRKAKDYGVVGFV
RAWEVTYGEANGWHPHWHVLFVFDKPLTPEOGDALEEVLYEAWSTALQDVGLYDFRE
HGVRLDLSHGEGGPRAYLMKYODGKAATTAEMTRDTKAGRDGHRTPPEIARVL
LTEDAADDRAQTVRLWQVETRAAGRMALYWSNGLRKLALBLDTRTGEITAAEE
RQGEALAVILADPWHQHIAARSGRLQLLKAEXGQDKVALVESWGLVMDVLPFP
PFVLEAAQRAEELDRRGVEGADQRKAHRVAKKSAVPAARKASAAAKTAPKKA
AHCTECGALDRLAGASGAGRHC"

```

ORIGIN

Alignment Scores:	7.63e-20	Length:	9367
Pred. No.:	452.50	Matches:	128
Score:	42.82%	Conservative:	54
Percent Similarity:	30.12%	Mismatches:	158
Best Local Similarity:	22.57%	Indels:	85
Query Match:	1	Gaps:	14
DB:			

US-10-007-527A-2 (1-379) x SNA243257 (1-9367)

Qy	19	ValSerSerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnIleThr	38
Db	6602	GTGCGGAGCATCGCGGATGGCGTCACGACTCCGTAGGGCGCTGTACGACGTGACG	6661
Qy	39	ThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIle	58
Db	6662	ACGACACAGCGCTCAGCGCTGCGCGGTACGCGCTCGCT---GGCGGGTCAACCC	6718
Qy	59	ValAsnGlyProLysGlySerGlyPhe---GlyGlyLeuArgSerCysGlyGlyTrp	77
Db	6719	AAGTGACCCCGCGCGAAGCGCTACCTCGCGGGCTGGCCACCTGGCGCAAGTGCAC	6778
Qy	78	IleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal	97
Db	6779	ATCTGCCCTGCTCGCGGCGAAGATCGGTCCGCCGTAGCTCGAATCCAGCGGCGC	6838
Qy	97	-----	97
Db	6839	GGGACGCGCTGGGAGGACATCGCGGCATCAACGCGCCGGAAGATCCGCCCGCGCGCC	6898
Qy	98	-----ValAla	99
Db	6899	GCCTGGAAGAAGACATCAACGACCGCAGCGGAGGACTGGGCGGCATCGAGAGGCC	6958
Qy	100	HisGlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGly	119
Db	6959	GATGGCTCCAGCGTGGTGGCTCGCATGCTGACGCTGACCATGGGCATCTCCCGG	7018
Qy	120	GlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsn---	138
Db	7019	CACACGCTCGCGAGCTGTGTACCCAGCAGCGGACGCGTGGGAAGGCGCTCGGCCG	7078
Qy	139	-----GlyArgArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg	156
Db	7079	AACGCGCGCGCGACCTGGCGCAAGCGAAGAGGACTAGCGCGTTCGCGCTTCGTCGG	7138
Qy	157	AlaValGluIleThrHisGlyLys---AsnGlyTrpHisValHisValHisAlaLeuLeu	175
Db	7139	GCCTGGAGGTCACTACCGCGAAGCAACGCTGGCACCCGCGCTGGCCTGCTCGTGTG	7198
Qy	176	MetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAsp	195
Db	7199	TTCTTCACACAGCGCTGACCCCGGAGCAGCGCGCTGGAGAGGTCTATCTACGAG	7258
Qy	196	ArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGly---	214
Db	7259	GCATGCTCCACCGCGCTTCAGGACGCTCGCGCTACCTCGCGCGACCGTGAACACGCG	7318
Qy	215	GlyLeuAspValArgLysIleGlyGlyGluAlaAspGlnValLeuAlaValLeuThr	234
Db	7319	CGCCTCGACCTGTCGGGCATGTTAGGCGGACCGCGG-----GCCCGGTACCTCATG	7372

```
QY 235 LysileAlaSerGly-----ValGlyMetGluValGlySerGlyAspGly 249
|||
Db 7373 AAGTACCAGGACGGGAAAGCGGCTGGACACCGCGCGCGAGATGACGCGGACGACG 7432
|||
QY 250 LysSerGlyArgHisGlyAsnArgAlaProTyrGluLeuLea-----Val 264
|||
Db 7433 AAGGCTGGCGCGACCGGCCACCGCCCGCTGACTGCTCCAAACGCGCTACGCAAGCGCTCGCGCGCCCTG 7492
|||
QY 265 AspAlaValGlyGlyAsp---ProGlnAlaLeuGluLeuTyrArgGluPheGluPheGly 283
|||
Db 7493 GATGCCCGGACGACGACCGTGGCCGACCGTGGCGGTGTGGCAGGAGTACGAGCGGCG 7552
|||
QY 284 SerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAlaGlyLeuGly 303
|||
Db 7553 GCGCGCGGATGCGCGCGCTGTACTGCTCCAAACGCGCTACGCAAGCGCTCGCGCGCCCTG 7612
|||
QY 304 AlaGluLeu-----ThrAspAlaGluLeuValGluGlnGluSerAlaProVal 320
|||
Db 7613 GTCGAACTCGACACCGCCGACCGGCGGAGATGCCCTGAGGAAACGCCAGGCGGCGCC 7672
|||
QY 321 MetValAlaIleIle-----ProAlaArgSer 329
|||
Db 7673 CTGGCGGTCTCTCGCGCGACCCCTGGCACCAGCACATGCGCCGCGCAAGGCGCGTCC 7732
|||
QY 330 TrpMetMetIleArgThrCysAlaProTyrValPheGlyGluLeuLeuGlyLeuValGlu 349
|||
Db 7733 CTCCAGCTGCTCAAGGCGCGCGAGAGCGCGCGCCGAGCAAGTCCGTGCGCTGGTGAG 7792
|||
QY 350 Ala---GlyAlaThrTrpGluAsnLeuArgAspHisLeuHisTyrArgLeuProAlaAla 368
|||
Db 7793 TCCGCGGCGCGTGGTGG-----GGAGCG 7816
|||
QY 369 AspValArgProPro 373
|||
Db 7817 GACGTCCTTCGCGCT 7831
|||

RESULT 12
AY172684/c 4603 bp DNA circular BCT 05-SEP-2003
LOCUS Corynebacterium glutamicum plasmid pAG3, complete sequence.
DEFINITION
ACCESSION AY172684
VERSION AY172684.1 GI:27657761
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 (bases 1 to 4603)
Tsuchi, A., Puhler, A., Kalinowski, J. and Thierbach, G.
Plasmids in Corynebacterium glutamicum and their molecular
classification by comparative genomics
J. Biotechnol. 104 (1-3), 27-40 (2003)
JOURNAL
MEDLINE 22830013
PUBMED 12948627
REFERENCE
2 (bases 1 to 4603)
Tsuchi, A., Puhler, A., Kalinowski, J. and Thierbach, G.
Direct Submission
AUTHORS
TITLE Submitted (05-NOV-2002) Department of Genetics, University of
JOURNAL Bielefeld, Universitaetstrasse 25, Bielefeld D-33615, Germany
FEATURES
source
1..4603
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/strain="22220"
/db_xref="taxon:1718"
/plasmid="pAG3"
complement(200..397)
/notes="orf1"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO18193.1"

CDS
/db_xref="GI:27657762"
/translation="MSKOKLSPPRAKVIAAVDREVPVQVAHKIGSTROKAGSALMAL
HKANLIDRFIRGIIYAPLEVDV"
complement(613..2061)
/gene="repa"
/gene="repa"
complement(613..2061)
/notes="replication protein; essential for pag3
replication"
/codon_start=1
/transl_table=11
/product="Repa"
/protein_id="AAO18194.1"
/db_xref="GI:27657763"
/translation="MSILASKQAEHATSPNQATPATHTTGLTGTDTTTHKHVKALAS
PALFRAKLMDLQRTMKVYEDTKLACGGRVAGGSVASLWMOAGRGARQGLQNSHS
WMSGVASSVITCRRANEVDAAIQAWANGAGLPAHYRGVSTSPSDEKQKRGVSNPVV
ERGISLMTITLRNSQSTEVWDATAGCQVAVTNTAAAGGATGADKSGRYGLAHY
RAIEVTHGKMGHVLHLVDFHDLVLSVDERDSLADRVFDRWAAKAVRLGWRAPSRDR
GIDVHVAASSDDAKSTGGYTCCKMLSGIAAETTTGGVTKKAGDNRNTPFOILGDLGK
QYTKRDHALWEKSGKRRGTGWSQGTGKVLGDELINELSDQIDDSLGDNDSQEVVAM
VGRGAWALASDTEKRAVILDAVAATDSAGAKAARDVLKLFVGAHTVTAIGLAKTD
LDIVLPHTDVFRELHEAVWSGRKQTALPVL"
2396..2785
/notes="orf3"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO18195.1"
/db_xref="GI:27657764"
/translation="MYRSATCALHALNVVRCFVGVSRETLDMTDYATLNRKICMAT
KAELOOLKDMQALKEQOOLAFNAREELEQORNLKPVVTVRVPOYVRDEMRKLAAL
CGWTLQQAANAFAAGMIDELSEQDQEH"
complement(2840..3232)
/notes="orf4"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO18196.1"
/db_xref="GI:27657765"
/translation="MTKPVNRFPSANDIEAMIGLVIAMNENYCTDKQAKTAKAIAIA
LNKWADELRKRGKDLAHDDETHIERVIRHLPFEQKATHDNQOLDEVHSLQSVHR
TEAAGEPTRYGHNVGVVOLLAEFDVE"
complement(3344..4030)
/notes="orf5"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO18197.1"
/db_xref="GI:27657766"
/translation="MATTPSKPFRGLAEQLSKVRSSLEVIASNVNADHDLSDGKNNK
NRYTPAPHTYIAQVEVALETISMSIDKAFNDARKALPTATDTSKLVVENELORII
NRGIPESVDGIYKLVTSHEPSTPTALIELEARGLSDSMVWSGLIKTSPEVAARTE
NMIOHARINSVFGYNLRTMYKALDRKAIENHWNVNTRSDADYDMEVPVTVFPVFPFK
TNAETVYRAS"
complement(4112..4519)
/gene="orf6"
complement(4112..4519)
/notes="orf6"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO18198.1"
/db_xref="GI:27657767"
/translation="MGRNATANLQVSTQIDNI SKALNIDWGGDHVAKANARASQL
QSAKLELADKPKRIASLELDSVLKDTCTPGERDQIMIAAISDKFLKLTSTGAPQ
EVAGQADAMAALDRFMYATKAAVHSGSPDNPNDH"
ORIGIN
```


Alignment Scores:

Pred. No.:	1.92e-18	Length:	4603
Score:	425.50	Matches:	109
Percent Similarity:	49.07%	Conservative:	75
Best Local Similarity:	29.07%	Mismatches:	138
Query Match:	21.22%	Indels:	53
DB:	1	Gaps:	14

US-10-007-527A-2 (1-379) x AY172684 (1-4603)

QY	2	ThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSerSer	21
Db	1971	ACCACCTGACCAACTAAGCACCATAAAGGTCAAGGCATATCACACGCGCTTTC	1918
QY	22	AspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGlu	41
Db	1917	-----CGCGCTAAACCTTGGGACTTACAGCGCCAGATGTGGAAGTCACAGAGGATAAA	1864
QY	42	ThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyVal-----ThrIleValAsn	60
Db	1863	ACCCTTCCCGCTGCGGTGCTTGGGTGGCGGTGGCTCCGCTGTTCATCGTTGGTGTGG	1804
QY	61	GlyProLysGly---SerGlyPheGlyGlyLeuArgSerCysGlyLysGlyIlePheCys	79
Db	1803	CAGCACAGGGGGCGGTGCACGCTGGGGTGTTCAGAACTCACACAGTGTTCGGGGTCG	1744
QY	80	ProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValVal---	98
Db	1743	CCTGTTGCATCGTCTGTGATCTGTTACGTGCTGTCTAATGAGGTGTGATGCAGCGATT	1684
QY	99	-----AlaHisGlnLeuGlyThrGlySer-----	106
Db	1683	CGGTGGGCTAATGGTGTGCTCTCACTCTCTGCTCATTAACGGGGCGTGTCACCTCC	1624
QY	107	-----ValAlaMet	109
Db	1623	AGTGATCGGAAAACAGCGCGCGGTGTCTGTGAATCCGGTGGTTGAACGGGGTATTAG	1564
QY	110	ValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeu	129
Db	1563	ATGACGTTGACGCTTAGGCAATACTCGACAGTCCCTAAGTGGGTGGGACGCTATT	1504
QY	130	SerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArg-----	143
Db	1503	GCTGGCTGTTGGCAGCGCTGTACTAAACACCGCGCGTGGCGTGGTCTCGCACTGCA	1444
QY	144	ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGly	163
Db	1443	GGCGATAAAGTCGCTATGGCATGTGCCATGTGGTACCGGGCTATTGAGGTCAACCCAG	1384
QY	164	LysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGlu	183
Db	1383	AGAATGGCTGGCATGTTTCACCTGCATGTGTGTGTTTCCATGATCGGGTGTGTAGTGT	1324
QY	184	AsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSer	203
Db	1323	GATGAAGAAGATCTTTGGCTCATCGTGTGTTTCGATAGTGGCTGCTAAGGCTGTTCGT	1264
QY	204	LeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGly	223
Db	1263	CTTGGTATCGGTGCACCGCTCTCGTATCGC-----GGAATTGATGTTGTGTCATGTT	1207
QY	224	GluAlaAspGlnVal-----LeuAlaAlaIleLeuThrLys-----IleAlaSerGlyVal	240
Db	1206	TCGTCTGATGACCTAAAAGCATTTGGTGGCTACCTGTAAAGGCGATGTTGTCCGATATT	1147
QY	241	GlyMetGluValGlySerGlyAsp---GlyLysSerGlyArgHisGlyAsnArgAlaPro	259
Db	1146	GCTCGGGAGACCAACACGGGCAGGTAACTAAGGAAGCTAAGGGAGATAACCGGACACT	1087
QY	260	TrpGluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGlu-----LeuTrp	277
Db	1086	TTTCAGATACTCGGTGAC---CTGGGAAAACAGTACACAAAGCGGTGATGCTCCCTATGG	1030

Qy	278	ArgGluPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArg	297
Db	1029	CTTGAGTGGGAAAAAGGTTCTAAAGGCGAGCGGCAGACTGGCTGGTTCACAAAGCATAAA	970
Qy	298	AlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGln-----Glu	315
Db	969	GATGTTCTTGGAAATTAAAC---GAACGTGCAGATGATCAGATTGATGATCTTTGGGTGAT	913
Qy	316	GluSerAlaProValMetValAlaIleIleProAlaArgSerTrp	330
Db	912	AACGATCAGTCCCGAAGTGGTGGCTATGGTGGCGCGAGCGGTGG	868

RESULT 13

AE017229	AE017229	300717 bp	DNA	linear	BCT 02-FEB-2004
LOCUS	Mycobacterium avium subsp. paratuberculosis str. k10, section 3 of 16 of the complete genome.				
DEFINITION	AE017229 AE016958				
ACCESSION	AE017229.1	GI:41395023			
VERSION					
KEYWORDS					
SOURCE	Mycobacterium avium subsp. paratuberculosis str. k10				
ORGANISM	Mycobacterium avium subsp. paratuberculosis str. k10				
	Bacteri; Actinobacteria; Actinobacteridae; Actinomycetales;				
	Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium				
	avium complex (MAC).				
REFERENCE	1. (bases 1 to 300717)				
AUTHORS	Li, L., Bannantine, J., Zhang, Q., Amonsin, A., Alt, D. and Kapur, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-SEP-2003) Biomedical Genomics Center, University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA				
FEATURES	Location/Qualifiers				
source	1..300717				

gene

gene

CDS

gene

Percent Similarity:	40.94%	Conservative:	58
Best local Similarity:	27.96%	Mismatches:	136
Query Match:	20.40%	Indels:	128
DB:	1	Gaps:	18
US-10-007-527A-2 (1-379) x AE017229 (1-300717)			
QY	22	AspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGlu	41
DB	275559	GAGCGCGCGGATTCGTTGGGGCGCAAGGGCGATCTGTGGCAGCGCTCTCGTTGAAG	275618
QY	42	ThrPheAsnAlaCysGlyArg	52
DB	275619	GCGGTGCGCTGCTGTGGCGGGTGTGCATAACGATCGGTTGGCGATCCTCATGACGCT	275678
QY	53	ValAenGlyValThrIle	68
DB	275679	---CAGGGCGTGGTATTAAGCGTCGGAGGTGATGAGCGGATGGTCGGAGTCTGCAT	275735
QY	69	GlyLeuArgSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAla	88
DB	275736	GGTCTGATGATCGGTTCCGTTCCGTTGGCGATSCCCCGGTTGTCGGCAGTCATGCGAT	275795
QY	89	HisArgAlaAspGluIleSerGlnVal	106
DB	275796	ACCAGGGCGCGGAATCGGTCCGACGTGAGAGAGTGTATCTGCTGGT---GGTGGC	275852
QY	107	ValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrp	126
DB	275853	GTGTATCTGTGATCGTTGACGATGCGGATGACCGGAGAGATGTTCTGTGACCTCTGG	275912
QY	127	ThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAenGlyArgArgTrpArgThrGluArg	146
DB	275913	GATTCGGCTTTCACGGCGCTGCGTTCGGTGTTCGGAACTCGTAATGGACGGGGCAGAA	275972
QY	146	-----	146
DB	275973	GAACGGATGTGACGGCAGCGGTGGCTTCCCGGAGATCATGGTGATGCC	276032
QY	147	GluMetTrpGlyCysAspGlyTrpValArgAlaValGluIleThrHisGlyLys	164
DB	276033	GAGCGTTTCACATAGCTGGGTCCACCGGTTGTGGAGCGACTTACGGCAAGCCGGAA	276092
QY	165	-----AsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal	181
DB	276093	CTTGGTGGCATGGCTGGCACTTCATATCCATCCCTCGTGTCTCAGTGACCAAGCTTG	276152
QY	182	SerGluAsnIleLeuGluSer	188
DB	276153	TCGAGCGCTCTGATTGAGGGCATTTGACGGACTCTGGGCCCGGAGTTAATCATGTTGG	276212
QY	189	-----PheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSer	203
DB	276213	TTGGCGCGCAACGCTTTGTCTGCTCGAATACATCAGCGCTGTGCGAGGATTTGCCAAG	276272
QY	204	LeuGlyPheAlaAlaProLeuArgAsnSerGlyLysLeuAspValArgLysIleGlyGly	223
DB	276273	GCTGGCTGTCCAGATGCCG-----GGGTCCGTCGGTTGATGTCGAGAGATTGACGAC	276326
QY	224	GluAlaAspGlnValLeuAlaAlaTyrluThrLys	239
DB	276327	GAGGCGCTGAATACGTGTGGCCGATACCTGTCCAAGGCTACATATGATGTCGGCGCAGC	276386
QY	240	ValGlyMetGluValGlySerGly	255
DB	276387	ATAGGCTTCAGGTGGAGCTGGGTTTCAACGAAACGCTCGGCGAGAACGTT	276440
QY	256	AsnArgAlaProTrpGluIle	268
DB	276441	AATCAGACTCCGTTCAAGTGTTCGAACTCTCGCTCAGTCGGTGGATGCTCGCGGGTTC	276500
QY	269	-----GlyAsp	270

Best Local Similarity:	27.23%	Mismatches:	169
Query Match:	19.93%	Indels:	77
DB:	6	Gaps:	15
US-10-007-527A-2 (1-379) x E17316 (1-2297)			
QY	4	ValSerAlaGluHisLeuSerGly---LysAspArgPro-----	15
DB	650	GTAGCGCGGAGAGCGGCAAGCGGCGGTAGCGCGCGCTTGGGAAAAGACGAATAAA	709
QY	16	-----ProValLeuValSerSerAspLysArg-----GlyIleArgHisGlu	29
DB	710	ATAACCCCTCCCTCTCCCGCACCGATTACCGCGCTTGGCGGTATGTCGCCGCGCTGAA	769
QY	30	LeuArgProLysLeuGlnGlnIleThrThrSerGluThrPheAsn-----	44
DB	770	AGCGGAAGATTCTCGTCCGTCATCGCGGTGGGAAACGCTCGGATTTCGAGCGGATTAAAG	829
QY	45	-----AlaCysGlyArgProLysSerGlyValAsnGlyValThrIleVal	59
DB	830	CTGCGCGGCTGCGCACGCTGCGGCAACCGGTGGAC-----ACCGCGGTGGGTGTCATG	883
QY	60	AsnGlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrIleCys	79
DB	884	ACCACGCGGAGAAAGCCGGTTTACAGGACCATGCTGCGGCTCGATCTGGGCAATGC	943
QY	80	ProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValala	99
DB	944	CCCACCTGCTCGGCAATCATTTGCCACGACGCGCCACGAGGTGCGCTCGCATCGGG	1003
QY	100	HisGlnLeu-----	102
DB	1004	AACCATGCGGAGAACTGAGAAAGCCCGCCGACCAATGGCAGGAGCAACATGAGGGG	1063
QY	103	-----GlyThrGlySerValAlaMetValThr	111
DB	1064	CAGCGTCTGCGCGCGAACTCATGTTGCCACAGCTTCGGAACTACATCTTCGGCAC	1123
QY	112	MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAla	131
DB	1124	CTCACCTGCGCCATGATCGGACAAATGCCGCTGGCCATGACCTTCGACCAATCTCAAG	1183
QY	132	AlaTyrLysAlaAlaThrAsnGlyArgArgTyrThrGluArgGluMetTyrGlyCys	151
DB	1184	GGCTGGCAGAAAGATGATTAAACGGAAGCCCTTGGCAACGCGGCTCGGAACGTGGAAATC	1243
QY	152	AspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGlyTyrHisValHisVal	171
DB	1244	AGGGTTTCGTCGCGCGGATTGAATACCTACGCTGTGAACGGCTGCACCCCTCACATT	1303
QY	172	HisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAsp	191
DB	1304	CATTTCGTGATGTTCTCGATGGCGATCTGGACGATGGCGAGCGGTGAGCAATGCAGCAA	1363
QY	192	AlaMetPheAspArgTyrThrSer-----LysLeuValSerLeuGlyPhe-----	206
DB	1364	TGGCTGTCGATCGCTGGAAACCATGTGCAAGCGGTGGCAAGCATACAGAAATAA	1423
QY	207	-----AlaAlaProLeuArgAsnSerGlyGlyLeuAspValArg---	219
DB	1424	GACGGCAACCCCTACACGTCGCCCG---AACGACGACACGCGCATAGATCTGCAATTC	1480
QY	220	LysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGly	239
DB	1481	AAATCGCGGCAAGGACGCGGAAACCGCTCGGCGCGAATACATCACCAAGATTCAAGCGCAC	1540
QY	240	ValGly-----MetGluValGlySerGlyAspGlyLysSerGlyArgHis	254
DB	1541	AAAGCGCGGTCATCTCGCTCAGGAATTCGCGCGCGGATATCAAGATGTCGTATG	1600
QY	255	GlyAsnArgAlaProTyrGluIleAlaValAspAlaValGlyGly-----AspProGln	272
DB	1601	GGGTGGTTAACCGGTTCCAAATTGTCGACTCCGGTGCCTCGGCGTCTCCGATTTCCAG	1660
QY	273	AlaLeuGluLeuTyrArgGluPheGluPheGlySerMetGlyArgAlaIleAlaTip	292
DB	1661	CGCGAAGATCTCTGGCTCGAATACTGGCAGCGCCACCTCTGCGCGCGCTGCATPAATGG	1720
QY	293	SerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleVal	312
DB	1721	TCGCTGGCTCAAGGAAGACATGGAGTTC---GAGGAACCTGGAAGACGAGGAGCTGGCG	1777
QY	313	GluGlnGluGluSerAlaProValMetValAla---IleIleProAlaArgSerTyrMet	331
DB	1778	GAGAAAGCCGACGAACCTGCGCGCTGCTGGTATGTCGCGCAATCGGTTTACAAA	1837
QY	332	MetIleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGluAlaGly	351
DB	1838	GACATTCGCAAGAGTGGCCTGAGACACTGGCGACGACATGGATGCCCGCGAACGC---	1894
QY	352	AlaThrTyrGluAsnLeu	357
DB	1895	GAAGACTGGCAGGAGTTC	1912
RESULT 15			
AF085719/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. 5750			
/organism="Bifidobacterium breve"			
/mol_type="genomic DNA"			
/strain="NCFB 2258"			
/db_xref="taxon:1685"			
/plasmid="pCIBb1"			
1463. .2062			
/gene="par"			
/note="partitioning gene"			
1463. .2062			
/gene="par"			
/note="par protein; putative involvement in plasmid			
partition"			
/codon_start=1			
/transl_table=11			
/product="partitioning protein"			
/protein_id="AAD34708.1"			
/db_xref="GI:4972587"			
/translation="MILSITASKGKTGKTSIFLACAAHTADPSLSITLADADPQOS			
ATDWAEEAQNKDPLFMVYKAGSDIENKNEGLTIIDTAPGFTDLTQILAKSDL			
CIITTSPTQIDRAEAWKTFAIQAGGTSAIITRANTRTKDFRMDIDLEHNDNANG			
ILFDTVIRQSTRYKSPFGHTPKTYDYADAWQEIQKAGE"			
complement(2373. .3515)			
/gene="rep"			
/note="replication gene"			
complement(2373. .3515)			
/gene="rep"			

Search completed: October 23, 2004, 05:46:07
Job time : 5476 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 03:58:12 ; Search time 641 Seconds

(without alignments)
3103.792 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005

Sequence: 1 MTSVSAHSLGKDRPPVLVS.....HLHYRLPAADVRRPIISVRK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10007527/runat_22102004_114115_26609/app_query.fasta_1.519
-DB=N_Geneseq_23Sep04 -QPM=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptlo -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10007527@cgn_1.1_470@runat_22102004_114115_26609 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2005	100.0	1140	6 ABQ76122	Rhodococc
C 2	2005	100.0	6334	6 ABQ76124	Rhodococc
C 3	2005	100.0	9652	6 ABQ76126	Plasmid p
C 4	2005	100.0	11241	6 ABQ76125	Plasmid p
C 5	2005	100.0	11241	12 ADH10182	Adh10182 E. coli-R
C 6	399.5	19.9	2297	2 AAV58945	AAV58945 B. breve

C	7	337	16.8	9431	2	AAQ20907	DNA encod
	8	326	16.3	4447	3	AAA90934	Aaa90934 B. lactof
	9	326	16.3	4447	6	ABL49733	Brevibact
	10	326	16.3	4447	6	AAD22582	B. lactof
	11	318	15.9	4447	3	AAA90951	B. lactof
	12	318	15.9	4447	6	ABL49734	Brevibact
	13	318	15.9	4447	6	AAD22583	B. lactof
	14	318	15.9	4447	10	ADB66212	B. lactof
C	15	296	14.8	8500	6	ABA93871	E. coli/c
C	16	282	14.1	2401	6	AAAS17124	Replicon
C	17	282	14.1	4005	6	AAAS17120	Ketogulon
	18	273	13.6	2534	3	AAZ39508	DNA seque
	19	273	13.6	3741	3	AAZ39509	DNA seque
C	20	214	10.7	5648	3	AAZ46134	Endogenou
C	21	133	6.6	349980	6	ABQ81845	Sequence
C	22	127	6.3	3331	1	AAAN91248	Seq
C	23	125	6.2	349980	6	ABQ81844	Bifidobac
C	24	125	6.2	349980	6	ABQ81844	Bifidobac
C	25	125	6.2	349980	6	ABQ81842	Bifidobac
C	26	125	6.2	349980	6	ABQ81842	Bifidobac
C	27	122.5	6.1	110000	4	AAI99682_02	Continuation (3 of
	28	122	6.1	1686	11	ABD01551	Pseudomon
C	29	122	6.1	1803	11	ABD01556	Pseudomon
C	30	122	6.1	2610	11	ABD01571	Pseudomon
C	31	122	6.1	2871	4	AAAS4190	Pseudomon
C	32	122	6.1	2871	8	ACA42440	Prokaryot
C	33	122	6.1	3790	10	ADC39175	Novel hum
C	34	122	6.1	5054	10	ADF74204	Human nov
C	35	122	6.1	5236	10	ADC30279	Human nov
C	36	122	6.1	5372	12	ADQ25367	Human nov
C	37	120.5	6.0	110000	4	AAI99683_02	Continuation (3 of
C	38	119	5.9	1428	11	ABD09766	Pseudomon
C	39	119	5.9	1910	2	AAV41758	Nitrosomo
C	40	119	5.9	2427	8	ACA37583	Prokaryot
C	41	119	5.9	3528	11	ABD09869	Pseudomon
C	42	117	5.9	135638	10	ABX34289	S. atrool
C	43	117	5.8	68750	3	AAZ55887	Sorangium
C	44	116	5.8	1542	6	AAD26199	Corynebac
C	45	116	5.8	4466	6	ABN59752	Novel hum

ALIGNMENTS

RESULT 1
ABQ76122
ID ABQ76122 standard; DNA; 1140 BP.
XX
AC ABQ76122;
XX
DT 13-JAN-2003 (first entry)
XX
DE Rhodococcus AN12 replication protein Rep DNA.
XX
KW Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase; gene;
KW shuttle vector; Rep; ds.
XX
OS Rhodococcus erythropolis.
XX
PN WO200255709-A2.
XX
PD 18-JUL-2002.
XX
PF 12-DEC-2001; 2001WO-US047868.
XX
PR 12-DEC-2000; 2000US-0254968P.
XX
PA (DUPO) DU FONT DE NEWMOURS & CO E. I.
XX
PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

DR WPI; 2002-557827/59.
DR P-PSDB; ABB84278.
XX
PT New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.
XX
PS Claim 2; Page 63-64; 96pp; English.

XX This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, an alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence encodes the Rhodococcus AN12 Rep protein
CC described in the disclosure of the invention

XX Sequence 1140 BP; 241 A; 287 C; 347 G; 265 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.41e-157 Length: 1140
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-007-527A-2 (1-379) x ABQ76122 (1-1140)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
DB 1 ATGACCAGCGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCGCTCTCGTGTG 60
QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnLeuThrThrSer 40
DB 61 TCCGATAAGCGCGGCATCCGGCAGCACTCGGACCCCAAACTTCAACAAATCACCACGTCA 120
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
DB 121 GAAACATTTAAACGCTGTGGCGCGCGAATTTCTGGCGTGAACGGGTGACCATTTGTCAAC 180
QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyLysGlyLysGlyPro 80
DB 181 GGTCCGAAAGGTCTCGAATTCGGAGCCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 240
QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
DB 241 TGCTGTGGGGAAGTCCGTGCACATCGTCAGACGAAATTTCTCAAGTTGTGTCTCAT 300
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
DB 301 CAATTCGGGACTTGGATCTGTTCGGATGTGTGACGATGACCATCGCCATACAGCTGTGTAG 360
QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
DB 361 CGGCTCCAGCACTATGACTGGACTTTTCGGCAGCCTCGMAAGCTCGCAACACGGTGTGT 420
QY 141 ArgTrpArgThrGluArgGluMetThrGlyCysAspGlyTyrValArgAlaValGluIle 160
DB 421 CGTTGGCGTACGGAACGTTGAAATGTACGCTCGCAGCGGATACGTGCGCTGTGAAATC 480
QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
DB 481 ACTCAGGAAACCGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTCAGTGTGTGAC 540
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
DB 541 GTGAGTGAAGAACATCTCGAATCTTCTCGGATGCGATGTTGATCGGTGGACTTCCAAA 600

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
DB 601 CTCGTATCTCTGGATTTGCTGCGCCACTACGTAATTCGGGTGGTCTCGATGTACGAAG 660
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
DB 661 ATCGCGCGTGAAGCTGATCAAGTTCTCGTGTGCTATCTGACGAAATTCATCTGGCGTT 720
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
DB 721 GGTATGAGGTTCGTAGTGGCGACGGAAGTGGTGCACATGGCAACCGTGACCCCTGG 780
QY 261 GluIleAlaValAspAlaValGlyLysAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
DB 781 GAAATCCTCTGTTGATGAGTGGCGGGGATCCACAGCGTTGGAACTGTGGCGAGAATTT 840
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAla 300
DB 841 GAGTTTGGTTCGATGGGACGTCGGCAATCGGTGGTCCGTCGATTCGTCGCCGAGCT 900
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
DB 901 GGTCTTGGGCGAGAACTAACAGATGCTTCAGATCGTTGAGCAGGAAGATCTGCCCGGTC 960
QY 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
DB 961 ATGTTTGGGATCATTCGGCGCGATCGTGGATGATGATTCGACTTGTGGCCTTACGTC 1020
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
DB 1021 TTGCGCGAGATCTCTCGACTCGTCAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 1080
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
DB 1081 TTGCATTTATCGATTGCCCGCAGCGGATGTGGCGCCCGCAATATATCGGTTCGCAAG 1137
RESULT 2
ABQ76124/c
ID ABQ76124 standard; DNA; 6334 BP.
XX
AC ABQ76124;
XX
DT 13-JAN-2003 (first entry)
XX
DE Rhodococcus AN12 derived plasmid pAN12 DNA.
XX
KW Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector; circular; ds.
XX
OS Rhodococcus erythropolis.
OS Synthetic.
XX
FN WO200255709-A2.
XX
PD 18-JUL-2002.
XX
PF 12-DEC-2001; 2001WO-US047868.
XX
PR 12-DEC-2000; 2000US-0254868P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
DR WPI; 2002-557827/59.
XX
PT New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.

XX Claim 19; Page 68-71; 96pp; English.
 XX This invention describes a novel nucleic acid encoding a replication
 CC protein or a plasmid stability protein. The product of the invention is
 CC useful for expression of nucleic acid such as genes encoding enzymes
 CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
 CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
 CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
 CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
 CC Actinomycetales bacteria. The replication protein or plasmid stability
 CC protein are useful in cloning and expression vectors and particularly in
 CC shuttle vectors for the expression of homologous and heterologous genes
 CC in *Rhodococcus* sp. This sequence represents the *Rhodococcus* AN12 derived
 CC plasmid pAN12 DNA described in the disclosure of the invention
 XX
 SQ Sequence 6334 BP; 1386 A; 1807 C; 1681 G; 1460 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.49e-156 Length: 6334
 Score: 2005.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-007-527A-2 (1-379) x ABQ76124 (1-6334)
 QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
 DB 3051 ATGACCGGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCGGCTCTCGTGTG 2992
 QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
 DB 2991 TCCGATAAGCGCGCATCCGCGCAAGACTCGACCCCAACTTCAACAATCACCACGTC 2932
 QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
 DB 2931 GAAACATTTAAGCGCTGTGGCGCGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTC 2872
 QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
 DB 2871 GTCCGGAAGGTTCTGGATTGGAGGCGCTTCGTTCTCGGAAAGGGTGGATCGCC 2812
 QY 81 CysCysAlaGlyLysValGlyValAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
 DB 2811 TGCTGTCGGGAAAGTCGTGTCACATCGTCACACGAAATTTCTCAAGTTGTTCAT 2752
 QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 DB 2751 CAATCGGACTGGATCTGTTGGATGGTGCAGTACCATGACCATCGCCATACAGCTGGTCAG 2692
 QY 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
 DB 2691 CGGCTCCACGACCTATGGATCGGACTTTCGGACGCTGGAAAGCTGGACCAACGGTCTG 2632
 QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 DB 2631 CGTTGGCGTACGGAACGTGAATGTACGGTGGCGAGTACGATCGCGCTGTGTGAATC 2572
 QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 DB 2571 ACTCAGGAAAAACGCTGGCAGCTCCAGTTTCAGCGCTACTCATGTTCACTGGTGAC 2512
 QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
 DB 2511 GTAGTGAGAACATCTCGAATCTTCTCGGATCGCATGTTCCATCGGTGACTTCCAAA 2452
 QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLeuAspValArgLys 220
 DB 2451 CTCGTATCTCTGGATTGCTGGCCACATACAGTAATTTCGGGTGGTCTCGATGTACGAAAG 2392
 QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240

DB 2391 ATCGCGGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCGATCTGGCT 2332
 QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyValAsnArgAlaProTrp 260
 DB 2331 GGTATGGAGTTGGTATGTGGCGACGGAAAAAGTGGTGCATGGCAACCGTGACCTGG 2272
 QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
 DB 2271 GAAATCGCTGTTGATGCGTGGCGGGGATCCACAAGCGTTTGAACCTGTGGCGAGATTT 2212
 QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
 DB 2211 GAGTTTGGTTGATGGGACGTCGGCAATCGCGTGGTCCGCTGGATTCGTCGCCGAGCT 2152
 QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 DB 2151 GGTCTTGGGGCAGAACTAACAGATCTCAGATCGTTGAGCAGAAAGAAATCTGCCCGGTC 2092
 QY 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
 DB 2091 ATGGTTGCGATCATTCGCGCGCATCGTGGATGATTCGACTTGTGGCTTACGTC 2032
 QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
 DB 2031 TTCGCGGAGATCCTCGGACTCGTGAAGCTCGCGGACTTCGGAAAAATCTTCGTGATCAC 1972
 QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
 DB 1971 TTGATTTATCATTTGCCGCGAGCGGATGTGCGCCCCCGGATAATATGATCGTCAAG 1915
 RESULT 3
 ABQ76126/c
 ID ABQ76126 standard; DNA; 9652 BP.
 XX
 AC ABQ76126;
 XX
 DT 13-JAN-2003 (first entry)
 XX
 DE Plasmid PRHBR171 DNA.
 XX
 KW Plasmid stability protein; replication protein; ethylene forming enzyme;
 KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
 KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
 KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
 KW shuttle vector; circular; ds.
 XX
 OS Synthetic.
 XX
 PN WO200255709-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 12-DEC-2001; 2001WO-US047868.
 XX
 PR 12-DEC-2000; 2000US-0254868P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Bramucci MG, Cheng Q, Kostichka KV, Tomb J;
 XX WPI; 2002-557827/59.
 DR
 PT New nucleic acid molecule encoding replication protein/plasmid stability
 PT protein, useful in cloning and expression vectors, particularly shuttle
 PT vectors for expression of heterologous genes in *Rhodococcus* species.
 XX
 PS Claim 27; Page 72; 96pp; English.
 XX
 CC This invention describes a novel nucleic acid encoding a replication
 CC protein or a plasmid stability protein. The product of the invention is
 CC useful for expression of nucleic acid such as genes encoding enzymes
 CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic

CC	acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC	hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC	dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC	actinomycetales bacteria. The replication protein or plasmid stability
CC	protein are useful in cloning and expression vectors and particularly in
CC	shuttle vectors for the expression of homologous and heterologous genes
CC	in Rhodococcus sp. This sequence represents the Plasmid pRHBR17 DNA
CC	described in the disclosure of the invention
XX	
SQ	Sequence 9652 BP; 2097 A; 2729 C; 2559 G; 2267 T; 0 U; 0 Other;
Alignment Scores:	
Pred. NO.:	5,81e-156 Length: 9652
Score:	2005.00 Matches: 379
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	6 Gaps:
US-10-007-527A-2 (1-379) x ABQ76126 (1-9652)	
QY	1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
Db	5052 ATGACACCGTAAGTGTGACACCTTCCGGCAAGACCGGCTCCGTCCTGTGTGC 4993
QY	21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
Db	4992 TCCGATAAGCGCGGCATCCGGCACGAATCGGCACCAAACTTCAACAAATCACCACGTCA 4933
QY	41 GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
Db	4932 GAAACATTAAACGCTGTGGCGCGCGATTTCTGGCGGAACGGGTGTGACCAATTGTCAAC 4873
QY	61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db	4872 GGTCCGAAGGTTCTGGATTCGAGCGCTTCGTTCTCGGAAGGGCTGGATTCGCC 4813
QY	81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValValAlaHis 100
Db	4812 TGCTGTGGGGAAAAAGTCGGTCACATCGTCGACAGCAAAATTTCTCAAGTTTGTCTCAT 4753
QY	101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db	4752 CAACTCGGACCTGGATCTCTGGATGGTGGATGATGATGATGATGATGATGATGATGAT 4693
QY	121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db	4692 CGGCTCCACGACCTATGGACTGGACTTTTCGGCAGCCTCGAAAGCTCGACCAACGGTCT 4633
QY	141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db	4632 CGTTGGCGTACGGAACGTAAGTGTACGGTTCGCGATACGATACGATACGATACGATAC 4573
QY	161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db	4572 ACTCAGGAAAAAAGCGCTGGCAGCTCCAGCTTCAGCGCTACTCATGTTTCAGTGGTGAC 4513
QY	181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db	4512 GTGAGTGAGACATCTCGAATCTCTCGATGCGATGTTGATGCGTGGACTTCCAAA 4453
QY	201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLysLeuAspValArgLys 220
Db	4452 CTCGTATCTCTGGGATTTGTCGCCCATCTACGTAATTCGGGTGGTCTCGATGTACGAAG 4393
QY	221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpThrLysIleAlaSerGlyVal 240
Db	4392 ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTGACGAAATTTGATCTGGGTT 4333
QY	241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyLysAsnArgAlaProTrp 260
Db	4332 GGTATGAGGTTGTGTAGTGGCGACGGAAAAAGTGGTCGATGCGCAACCGTGCACCCCTGG 4273
QY	261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db	4272 GAATCGCTGTTGATGCAGTGGCGGGATCCACACGCTTGGAACTGTGGCGAGAATTT 4213
QY	281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db	4212 GAGTTTGTTGATCGGACGCTCGCGCAATCGCGTGGTCCGTTGATTTGCGTCCCGAGCT 4153
QY	301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db	4152 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGTTGAGCGGAGAAATCTTCGTGATCAC 4093
QY	321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
Db	4092 ATGGTTGGGATCATTCGCGCGGATCGTGGATGATGATTCGACTTGTGCGCCTTACGTC 4033
QY	341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db	4032 TTCGGCGAGATCCTCGGACTCGTCAAGCTGCGCGACTTGGGAAAAATCTTCGTGATCAC 3973
QY	361 LeuHisTyrArgLeuProAlaAlaAspValArgProIleIleSerValArgLys 379
Db	3972 TTGCATTATCATGATTCGCCGCGAGCGATGTGCGGCCCGCATATATCGTTTCGCAAG 3916
RESULT 4	
ABQ76125/c	
ID	ABQ76125 standard; DNA; 11241 BP.
XX	
AC	ABQ76125;
XX	
DT	13-JAN-2003 (first entry)
XX	
DE	Plasmid pRHBR17 DNA.
XX	
KW	Plasmid stability protein; replication protein; ethylene forming enzyme;
KW	carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW	polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW	alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW	shuttle vector; circular; ds.
XX	
OS	Synthetic.
XX	
PN	WO200255709-A2.
XX	
PD	18-JUL-2002.
XX	
PF	12-DEC-2001; 2001WO-US047869.
XX	
PR	12-DEC-2000; 2000US-0254868P.
XX	
PA	(DUPO) DU PONT DE NEMOURS & CO E I.
XX	
PI	Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX	
DR	WPI; 2002-557827/59.
XX	
XX	New nucleic acid molecule encoding replication protein/plasmid stability
PT	protein, useful in cloning and expression vectors, particularly shuttle
PT	vectors for expression of heterologous genes in Rhodococcus species.
XX	
PS	Claim 26; Page 71-72; 96pp; English.
XX	
CC	This invention describes a novel nucleic acid encoding a replication
CC	protein or a plasmid stability protein. The product of the invention is
CC	useful for expression of nucleic acid such as genes encoding enzymes
CC	involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC	acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC	hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC	dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC	actinomycetales bacteria. The replication protein or plasmid stability
CC	protein are useful in cloning and expression vectors and particularly in
CC	shuttle vectors for the expression of homologous and heterologous genes
CC	in Rhodococcus sp. This sequence represents the Plasmid pRHBR17 DNA

CC described in the disclosure of the invention

XX Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.98e-156 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-007-527A-2 (1-379) x ABQ76125 (1-11241)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
DB 6641 ATGACCAACGCGTAAGTGTGAACACCTTTCCGGCAAGACCGCGCTCCCGTCTCGTGTGC 6582

QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
DB 6581 TCCGATTAAGCGCGCATCCGGCAGCAACTGGGACCCCAAACTTCAACAAATCACCACGTC 6522

QY 41 GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
DB 6521 GAAACATTTAAACGCTGTGGCCGCGCATTTCTGGCGTGAACGCTGTGACCATTTGTCAAC 6462

QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80
DB 6461 GGTCCGAAGAGTTCTGGATTTCGGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCC 6402

QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
DB 6401 TGGTGTGGGAAAGTCGGTGCAATCGTGCAACGAAATTTCTCAAGTTGTTCCTCAT 6342

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
DB 6341 CAACTCGGGACTGGATCTGTTGGATGCGTACGATGACCATCGCCATACAGCTGGTCAG 6282

QY 121 ArgLeuHisAspLeuThrGlyLeuSerAlaAlaTyrLysAlaAlaThrAsnGlyArg 140
DB 6281 CGGCTCCACGACCTATGGACTGGACTGTTCCGACGCTGGAAAGCTGGACCAACGGTCTG 6222

QY 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
DB 6221 CGTTGGGTACGAGACGTAATGATGACGTCGACGATGACGTCGCGCTGTGTAATC 6162

QY 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
DB 6161 ACTCACGAAAAACGCGCTGGCAGCTCCAGCTTCAACGCGCTACTCATGTTCAAGTGTGAC 6102

QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
DB 6101 GTGAGTGAGAACATCTCGAATCTCTCGGATCGCATGTTCCGATCGGCTGCTCCAA 6042

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
DB 6041 CTGATCTCTGGGATTTGTCGCCACTAGTAATTCGGGTGCTTCGATGACGAAAG 5982

QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
DB 5981 ATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTGACGAAAAATTCATCTGGCGTT 5922

QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
DB 5921 GGTATGAGGTGGTATGTCGCGACGGAAGAAAGTGTGCATGACCAACCGTGCACCTCG 5862

QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
DB 5861 GAAATCGCTGTTGATGACGATGGCGCGGATCCACACGCTTGGAACTGTGGCGAATTT 5802

QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
DB 5801 GAGTTTGGTTCGATGGACGCTCGGGCAATCGCGTGTCCGCTGGATTGGTGGCCGAGCT 5742

QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
DB 5741 GGTCTGGGGCAGAACTAAACAGATGCTCAGATCTGTGAGCAGGAAGAAATCTGCCCGGTC 5682

QY 321 MetValAlaIleIleProAlaAspSerTyrMetMetIleArgThrCysAlaProTyrVal 340
DB 5681 ATGTTTGGGATCATTCGCGCGCGATCGTGGATGATGATTCGGACTTGTGCGCCTTACGTC 5622

QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTyrGluAsnLeuArgAspHis 360
DB 5621 TTCGGCGAGATCTTCGGACTCGTCGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562

QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProIleIleSerValArgLys 379
DB 5561 TTGCATTATGATTCGCCGACGATGTGGCGCCCGCATATATATCGTTCCGAAG 5505

RESULT 5
ADH10182/C
ID ADH10182 standard; DNA; 11241 BP.
XX
AC ADH10182;
XX
DT 11-MAR-2004 (first entry)
XX
DE E. coli-Rhodococcus shuttle plasmid pRHBR17.
XX
KW Aryl carotenoid; beta-ionone; carotene desaturase; CrtU; isorenieratene;
KW chlorobactene; pharmaceutical; food supplement; animal feed additive;
KW food colourant; cosmetic; crtO; ds.
XX
OS Synthetic.
XX
FN WO2003093200-A2.
XX
PD 13-NOV-2003.
XX
PF 06-MAY-2003; 2003WO-US014111.
XX
PR 06-MAY-2002; 2002US-0378312P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Cheng Q, Tao L, Rouviere PE;
XX
XX WPI; 2004-081862/08.
XX
XX Production of aryl carotenoids useful in e.g. pharmaceuticals and food
XX colorants involves bioconversion of cyclic carotenoids comprising beta-
XX ionone ring in the presence of carotene desaturase.
XX
XX Example 3; SEQ ID NO 13; 66pp; English.
XX
XX The invention relates to production of aryl carotenoid compounds. The
XX method involves transforming a host cell comprising a cyclic carotenoid
XX having at least one beta-ionone ring with a gene encoding a carotene
XX desaturase (CrtU) followed by growing the transformed host cell. Also
XX provided is a method for regulating aryl carotenoid biosynthesis in a
XX host cell. The method is useful in the production of aryl carotenoids
XX such as isorenieratene and chlorobactene, which are useful in
XX pharmaceuticals, food supplements and chlorobactene, which are useful in
XX pharmaceuticals, food supplements, animal feed additives, food colorants
XX and cosmetics. The present sequence represents a E. coli-Rhodococcus
XX shuttle plasmid pRHBR17 nucleotide sequence
XX
XX Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.98e-156 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-007-527A-2 (1-379) x ADH10182 (1-11241)

Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
 |||||
 Db 6641 ATGACCAAGCGTAAAGTGTGAACACCTTCCGGCAAGACCGGCTCCGCTCGTGTGCG 6582
 |||||
 Qy 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
 |||||
 Db 6581 TCCGATAAGCGCGCATCCGGCAGCACTGCGGACCAACCTTCAACAAATCACCACGTCAC 6522
 |||||
 Qy 41 GluThrPheAenAlaCysGlyArgProLysSerGlyValAenGlyValThrIleValAsn 60
 |||||
 Db 6521 GAAACATTTAACGGCTGTGGCCGCGGATTTCTGCGTGAACCGGTGACCATTTGTCAAC 6462
 |||||
 Qy 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
 |||||
 Db 6461 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCTGCTCGGAAAGGCTGATCTGCCCC 6402
 |||||
 Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
 |||||
 Db 6401 TGCTGTGCGGAAAGTCCGTGCATCGTCAGACGAAATTTCTCAAGTTGTGCTCAT 6342
 |||||
 Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 |||||
 Db 6341 CAATCCGGGACTGGATCTGTTGCCGATGTGACATGACCATGCCCATACAGCTGGTCAG 6282
 |||||
 Qy 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140
 |||||
 Db 6281 CGGCTCCACGACCTATGACCTGGACTTTCCGCGAGCTTGAAAGCTGCGACCAACGGTCGT 6222
 |||||
 Qy 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 |||||
 Db 6221 CGTTGGCGTACGGAACGTGAATGTACGGCTCGCAGCGATACGTGCGCGCTGTTGAAATC 6162
 |||||
 Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 |||||
 Db 6161 ACTCAGGAAAGAACGGCTGCAGCTCCAGCTTACGGCGTACTCATGTTGAGTGGTGAC 6102
 |||||
 Qy 181 ValSerGluAenIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
 |||||
 Db 6101 GTGAGTGAGAACATCTCCGATCTCTCGATGCGATGATGTTGATCGGTGGACTTCCAAA 6042
 |||||
 Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
 |||||
 Db 6041 CTCGTATCTCTGGGATTTGCTGCGCCACTAGCTAATTCGGGTGGTCTCGATGTACGAAAG 5982
 |||||
 Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 |||||
 Db 5981 ATCGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTTGCATCTGGCGTT 5922
 |||||
 Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
 |||||
 Db 5921 GGTATGGAGTTGTTAGTGGCGCAGGAAAGTGGTGCACATGGCAACCGTGCAACCTGG 5862
 |||||
 Qy 261 GluIleAlaValAspAlaValGlyClyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
 |||||
 Db 5861 GAAATCGCTGTGTGATGCTGGGCGGGATCCACAAAGCTTGGAACTGTGGCGAGAAATTT 5802
 |||||
 Qy 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
 |||||
 Db 5801 GAGTTTGGTTCGATGGGACGTGCGCAATCCGGTGTCCCGTGGATTCGTTGCGCGAGCT 5742
 |||||
 Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 |||||
 Db 5741 GGTCTTGGGGCAGAACTAACAGATCTCAGATCGTTGAGCAGGAAAGAAATCTGCCCGGTC 5682
 |||||
 Qy 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
 |||||
 Db 5681 ARGTTTGGATCATTTCCGGCGGATCGTGGATGATTCGGACTTGTGGCTTTACGCTC 5622
 |||||
 Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
 |||||

Db 5621 TTCGGCGAGATCCTCGGACTCGTCAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562
 Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
 |||||
 Db 5561 TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCGCCCGCATATATCGTTTCGCAAG 5505
 |||||
 RESULT 6
 AAV58945
 ID AAV58945 standard; DNA; 2297 BP.
 XX
 AC AAV58945;
 XX
 DT 05-JAN-1999 (first entry)
 XX
 DE B. breve essential region gene.
 XX
 KW Essential region gene; shuttle vector; ds.
 XX
 OS Bifidobacterium breve.
 XX
 FH Key Location/Qualifiers
 FT -35_signal 387..392 /*tag= a
 FT -10_signal 410..415 /*tag= b
 FT RBS 525..530 /*tag= c
 FT CDS 539..2023 /*tag= d
 FT
 XX JP10262670-A.
 XX
 PD 06-OCT-1998.
 XX
 PF 27-MAR-1997; 97JP-00091387.
 XX
 PR 27-MAR-1997; 97JP-00091387.
 XX
 PA (HONS) YAKULT HONSHA KK.
 XX
 DR WPI; 1998-587288/50.
 DR P-PSDB; AAW73071.
 XX
 PT Shuttle vector for a Bifidobacterium species - contains specific elements
 PT from existing plasmids e.g. pNBb1 of ATCC 15698.
 XX
 PS Disclosure; Page 14-15; 17pp; Japanese.
 XX
 CC This sequence represents the essential region gene from plasmid pNBb1
 CC used in the vector of the invention. The vector is a shuttle vector for a
 CC Bifidobacterium, and contains: (a) a replicated essential region
 CC originated from a plasmid pNBb1 of Bifidobacterium breve ATCC 15698; (b)
 CC a replicated essential region originated from a plasmid of E. coli; and
 CC (c) an antibiotic-resistant gene originated from a plasmid of E. coli and
 CC an antibiotic-resistant gene functioning by a Bifidobacterium. The
 CC shuttle vector can be used to transform various Bifidobacteria species
 XX
 SQ Sequence 2297 BP; 524 A; 659 C; 689 G; 425 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,97e-23 Length: 2297
 Score: 399.50 Matches: 116
 Percent Similarity: 42.25% Conservative: 64
 Best Local Similarity: 27.23% Mismatches: 169
 Query Match: 19.93% Indels: 77
 DB: 2 Gaps: 15

US-10-007-527A-2 (1-379) x AAV58945 (1-2297)

Qy 4 ValSerAlaGluHisLeuSerGly---LysAspArgPro----- 15
 |||||
 Db 650 GTAGCCGCGAAGCGGCAAGCGGCGCGCTTGGGAAAGACGAAATAA 709

Qy	228	ValLeuAlaAlaTyrLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly	245
Db	2913	GACCTTCGGCGATACATCGCAAGACCCAGACGGGAGCGCCCTCGAACTGCC	2854
Qy	246	SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrrpGluIle-----	262
Db	2853	CGCGCCGACCTCAAGACGGCGACCGCGGGGAACTCGCCCCGTTCGAACTCCTCGGACGG	2794
Qy	263	AlaValAspAlaValGlyGly-----AspProGlnAla-----	273
Db	2793	ATCCGGGACCTGACCGCGCGCATGACCGAGGAGCAGCCCGCGGGTCTGGCTCGCTGGAG	2734
Qy	274	-----LeuGluLeuTrrpArgGluPheGluPheGlySerMetGlyArgArgAlaIleAla	291
Db	2733	TGGAACCTCTCGCGCTGTGCACGAGTAGTACGACGGCGGCAACCCCGGGGACCCGGGGCCTCGAA	2674
Qy	292	TrrpSerArgGlyLeuArgAlaAa-gAlaGlyLeuGlyAlaGluLeuThrAspAla	309
Db	2673	TGACACCGCTACTCTGGCGGACAGAGTCTCGGCTCGACGGCGGGACACCGAGGCC	2620

RESULT 8	
AAA90934	
ID	AAA90934 standard; DNA; 4447 BP.
XX	
AC	AAA90934;
XX	
DT	15-SEP-2003 (revised)
DT	15-JAN-2001 (first entry)
XX	
DE	B. lactofermentum p48K coding sequence.
XX	
KW	Temperature sensitive plasmid; TSRCR; protein production;
KW	temperature sensitive replication control region; p48K; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PH	Key Location/Qualifiers
FT	CDS 1318..2601
FT	/*tag= a
XX	
XX	EPI038966-A1.
PN	
XX	
PD	27-SEP-2000.
XX	
PF	16-MAR-2000; 2000EP-00105326.
XX	
PR	16-MAR-1999; 99JP-00069896.
XX	
PA	(AJIN) AJINOMOTO CO INC.
XX	
PI	Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;
XX	
DR	WPI; 2000-573832/54.
DR	P-FSDB; AAY97536.
XX	
PT	Plasmids containing a temperature sensitive replication control regions
PT	useful for breeding microorganisms for the production of amino acids by
PT	fermentation.
XX	
FS	Claim 7; Page 14-17; 29pp; English.
XX	
CC	This sequence encodes the Brevibacterium lactofermentum p48K protein. The
CC	invention relates to a plasmid containing a temperature sensitive
CC	replication control region (TSRCR) and a marker gene (MG). The TSRCR is
CC	derived from plasmid pM330 harboured by Brevibacterium lactofermentum
CC	(ATCC 13869) and allows the plasmid to replicate autonomously at a low
CC	temperature but does not allow the plasmid to replicate autonomously at
CC	an elevated temperature in coryneform bacteria within a temperature range
CC	in which the bacteria can grow. The plasmid can be used for modifying a
CC	chromosomal gene in a coryneform bacterium, which may be used for the
CC	production of useful substances, such as amino acids, by fermentation, to
CC	change their genetic traits. Therefore, the plasmid can be used for

Alignment Scores:		5,53e-17	4447
Pred. No.:	Score:	326.00	101
Percent Similarity:	Conservative:	47.75%	58
Best Local Similarity:	Mismatches:	30.33%	133
Query Match:	Indels:	16.26%	42
DB:	Gaps:	6	14
US-10-007-527A-2 (1-379) x ABL49733 (1-4447)			
Qy	34	LeuGlnGlnIleThrSerGlyArgProIleSerGlyVal	53
Db	1396	ATGTACAAGATACCAATATAGAGCTCTCGCGGGTGCCATAGTG-GCGCAGGACGA	1454
Qy	54	Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr	71
Db	1455	AGCTGTTCGGTGTCTCGTTCGTCCTAACGGT-----GCTTCGCGAGTTTCAGGGTCTGCA	1508
Qy	71	gSerCysGlyLysGlyTripleCysProCysCysAlaGlyValGlyAlaHisArgAl	91
Db	1509	AAACTCTCACTCTCGCTGGGGGTCACTCTGGCTGAATTCGAATTCGCGCGAACCCG	1568
Qy	91	aAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetValth	111
Db	1569	CATTGAGCTGGCTATTGCTACTAAGAAATCACTTGGCGGGGTGGCGGCTCATGATTT	1628
Qy	111	rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe	130
Db	1629	TGTGGCAGCTGTCGACACACCGCTCACAGTCATTTTGGCAGGTTCGAAGCGGGTATTA	1688
Qy	130	rAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTygl	150
Db	1689	GACTGCGTACTCTTCGATGGTGGTGAACATCTCAGTGAAGAAAGAACGTCACCGTACGG	1748
Qy	150	yCysAspGlyTyrrValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa	169
Db	1749	GCTGAGACACCTATAGTACTATAGGTCACAGACTCTTGGCGGACGGTTCGCACCTT	1808
Qy	169	lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh	189
Db	1809	GCACCGCACATGCTGTGTTCTTGGATCGTCCACTGTCTGACGATGAACCTCAAGGCGTT	1868
Qy	189	eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr	209
Db	1869	TGAGGATTCATGTTTCCCGCTGCTGCTGCTGTTAAAGCCCGGTATGGACGCGCC	1928
Qy	209	oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGlyGluAlaAspGl	227
Db	1929	ACTGGTGAGCAGCGGGTCAAACTTGATCAGGTGTCTACTGGGTGGAGCGTGGCAA	1988
Qy	227	nValLeuAlaAlaTyrrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl	247
Db	1989	A---ATGGCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGGCTC	2036
Qy	247	yAspClyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa	267
Db	2037	CGTACTAAACCCGCTCTAAGGGGTGCTACACGCGCTTCAGATG---TTGGATATGTT	2093
Qy	267	lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl	279
Db	2094	GGCCGATCAAGACGACGCGCGGAGATATGACGCTGTTTGGTGGCTCGGTGGCTGA	2153
Qy	279	uPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr	299
Db	2154	GTATCAGGTGGTGTCTTAAACAACTCGGTTCG---TCCTGGTCACTGGG-----GCTAA	2204
Qy	299	gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu--	315
Db	2205	GGTGTCTGGGCATTGATTACATAGACGCTGATGCTGTGAATGGAAAGAAAGAACT	2264
Qy	316	-----GluSerAlaProValMetValAl	323

Db 2265 GTACAAAGCTCGCGGTCTCGGAAGCACCGGAACGGGTGCAATCAACCCGCTGCTCTTTC 2324

Qy 323 alleleProAlaArgSerTrpMetMetIleArgThr 335

Db 2325 TTTGGTGAAGCCCGATGATTTGGAACTGATTCAGTCT 2361

RESULT 10

AA022582

ID AAD22582 standard; DNA; 4447 BP.

XX AAD22582;

XX 29-AUG-2003 (revised)

DT 12-FEB-2002 (first entry)

XX B. lactofermentum pSF6 replication control region DNA.

Coryneform bacteria; arginine repressor; L-arginine; liver function;

amino acid infusion; ds.

XX Corynebacterium glutamicum.

OS

Key Location/Qualifiers

FT 1318..2598

CDS /*tag= a

FT /product= "B. lactofermentum pSF6 replication control

FT region Protein"

XX

EP1154020-A2.

PN 14-NOV-2001.

XX

24-APR-2001; 2001EP-00109457.

PF

28-APR-2000; 2000JP-00129167.

PR

(AJIN) AJINOMOTO KK.

XX

Suga M, Asakura Y, Mori Y, Ito H, Kurahashi O;

XX

WPI; 2002-012660/02.

DR

P-PSDB; AAE13530.

XX

Coryneform bacterium with inactivated arginine repressor, useful for

fermentative production of arginine at high levels.

XX

Example 1; Page 12-15; 30pp; English.

PS

XX

The present invention relates to Coryneform bacterium in which the

arginine repressor does not function normally and which produces L-

arginine. Coryneform bacterium is used to produce L-arginine which is

useful in agents for promoting liver function and in amino acid infusions

or in 'comprehensive' amino acid pharmaceutical. Coryneform bacterium

accumulate L-arginine to higher levels than wild-type strains, e.g. 120

compared with 20, mg/dl. The present sequence is Brevibacterium

lactofermentum pSF6 replication control region DNA. This DNA is used in

the construction of shuttle vector for Escherichia coli and Coryneform

bacteria and temperature sensitive vector. (Updated on 29-AUG-2003 to

CC standardise OS field)

XX

Sequence 4447 BP; 824 A; 997 C; 1378 G; 1248 T; 0 U; 0 Other;

SX

Alignment Scores:

Pred. No.: 5,53e-17 Length: 4447

Score: 326.00 Matches: 101

Percent Similarity: 47.75% Conservative: 58

Best Local Similarity: 30.33% Mismatches: 133

Query Match: 16.26% Indels: 42

DB: Gaps: 6

US-10-007-527A-2 (1-379) x AAD22582 (1-4447)

Qy 34 LeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53


```

Db 1396 ATGTACAAGATCACCATAAGTCTCTGGGGGTGCCTAGGTG-GCGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr 71
Db 1455 AGCTTTGGGGTGTCTGGTCTGCTAAACGGT-----GCTTCGAGTTTGGGGTCTGCA 1508
QY 71 gSerCysGlyLysGlyTyrPilleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AAACCTCTCACTCTCGCTGGGGTCACTCTGCTGAATGGAGTATGGGGGAACGGCG 1568
QY 91 aAspGluLeuSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATTGAGCTGGCTATTGCTACTAAGATCACTTGGCGGGGTGGCGGCTCATGATGTT 1628
QY 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisLeuPheLeuThrGlyLeuSe 130
Db 1629 TGTGGGCACTGTTCGACACACCGCTCACGTCATTTGCGCAGGTGAAGCGGGTATTAA 1688
QY 130 rAlaAlaTyrLysAlaAlaThrAsnGlyArgArgTyrArgThrGluArgGluMetTyrGl 150
Db 1689 GACTCGGTACTCTTCGATGGTGAAACATCTCAGTGGAGAAAGAACCTGCACGGTAGCG 1748
QY 150 YCysAspGlyTyrValArgAlaValGluIleThrHisGly---LysAsnGlyTyrHisVa 169
Db 1749 GGTGGGACACACCTATAGTACTATAGCTCACAGACTCTTGGCGGAACGGTGGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnLeuLeuGluSerPh 189
Db 1809 GCACCGCAACATGCTGTGTCTGGATCGTCCACTGTCTGACGATGAACCTCAAGCGGTT 1868
QY 189 eSerAspAlaMetPheAspArgTyrThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCCATGTTTCCCGCTGCTGTGCTGGTGAAGCCGGTATGACGGCGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGluAlaAspGl 227
Db 1929 ACTCGTGAGCAGCGGGTCAAACTTGATCAGGTGTCTACCTGGGTGGAGCGTGGCAA 1988
QY 227 nValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCAACCTTACTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGGCTC 2036
QY 247 YAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrPilleAlaValAspAlaVa 267
Db 2037 CGCTACTAAAACCGCGTCAAGGGTGTACACCGCTTTCAGATG---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTyrArgGl 279
Db 2094 GGCCGATCAAAAGCGACCGCGGAGGATATGGCGCTGTTTGTGGCTCGGTGGCTGA 2153
QY 279 uPheGluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaR 299
Db 2154 GTAAGAGTTGGTCTATAAAACCTGCGTTCG---TCCTGGTCACTGGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu-- 315
Db 2205 GCGTGTCTTGGCATTGATTACATAGACGCTGATGATGCTGCTGAATGGAAGAACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGGTCTGGAAGCAGCGGAACGGTGCATCAACCCGGTCTGCTTGC 2324
QY 323 alIleIleProAlaArgSerTyrMetMetIleArgThr 335
Db 2325 TTTGGTGAAGCCGATGATTGGAACACTGATTTCAGTCT 2361

```

RESULT 11

AAA90951

ID AAA90951 standard; DNA; 4447 BP.

XX

AC AAA90951;

XX

DT 15-SEP-2003 (revised)
 DT 15-JAN-2001 (first entry)

XX B. lactofermentum pSPK6 coding sequence.

XX Temperature sensitive plasmid; TSSCR; protein production;
 XX temperature sensitive replication control region; pSPK6; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers
 CDS 1318..2601
 /*tag= a

XX EPI038966-Al.

XX 27-SEP-2000.

XX 16-MAR-2000; 2000EP-00105326.

XX 16-MAR-1999; 99JP-00069896.

XX (AJIN) AJINOMOTO CO INC.

XX Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;

XX WPI; 2000-573832/54.

XX P-PSDB; AAY97537.

XX Plasmids containing a temperature sensitive replication control regions
 useful for breeding microorganisms for the production of amino acids by
 fermentation.

XX Disclosure; Page 19-22; 29pp; English.

XX This sequence encodes the Brevibacterium lactofermentum pSPK6 protein.
 The invention relates to a plasmid containing a temperature sensitive
 replication control region (TSSCR) and a marker gene (MG). The TSSCR is
 derived from plasmid pAM330 harboured by Brevibacterium lactofermentum
 (ATCC 13869) and allows the plasmid to replicate autonomously at a low
 temperature but does not allow the plasmid to replicate autonomously at
 an elevated temperature in coryneform bacteria within a temperature range
 in which the bacteria can grow. The plasmid can be used for modifying a
 chromosomal gene in a coryneform bacterium, which may be used for the
 production of useful substances, such as amino acids, by fermentation, to
 change their genetic traits. Therefore, the plasmid can be used for
 breeding microorganisms for the production of amino acids by
 fermentation. The plasmid comprises a TSSCR that allows the plasmid to
 replicate autonomously at an elevated temperature in coryneform bacteria
 within a temperature range in which the bacteria can grow. (Updated on 15
 -SEP-2003 to standardise OS field)

XX SQ Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,55e-16	Length:	4447
Score:	318.00	Matches:	100
Percent Similarity:	47.45%	Conservative:	58
Best Local Similarity:	30.03%	Mismatches:	134
Query Match:	15.86%	Indels:	42
DB:	3	Gaps:	14

US-10-007-527A-2 (1-379) x AAA90951 (1-4447)

QY 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53

Db 1396 ATGTACAAGATCACCATAAGTCTCTGGGGGTGCCTAGGTG-GCGCAGGACGA 1454

QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr 71

Db 1455 AGCTTTGGGGTGTCTGGTCTGCTAAACGGT-----GCTTCGAGTTTGGGGTCTGCA 1508

QY 71 gSerCysGlyLysGlyTyrPilleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91

OS XX Corynebacterium glutamicum.
FH Key Location/Qualifiers
FT CDS 1318..2601
FT /*tag= a
FT /product= "p48K"
XX
PN JP2002017362-A.
XX
XX 22-JAN-2002.
XX
XX 05-JUL-2000; 2000JP-00204236.
XX
XX 05-JUL-2000; 2000JP-00204236.
XX
XX (AJIN) AJINOMOTO KK.
XX
XX WPI; 2002-263243/31.
XX P-PSDB; ABB06342.
XX
XX Transformation of high temperature-resistant coryneform microbe and transformant.
XX
XX Example; Page 20-22; 28pp; Japanese.
XX
XX The present invention describes a method for transforming a high temperature-resistant coryneform microbe in which the high temperature-resistant coryneform microbe is treated with a chemical changing the structure of cell wall of the microbe and an electric pulse is applied to a solution containing the microbe having a changed structure of cell wall and a DNA. Also described are: (1) transforming a high temperature-resistant coryneform microbe in which the cell wall of a high temperature-resistant coryneform microbe is weakened and electric pulse is applied to a solution containing the microbe having weakened cell wall and a DNA; (2) a transformant of a high temperature-resistant coryneform microbe in which a recombinant DNA is introduced; and (3) a transformant of a high temperature-resistant coryneform microbe carrying both a plasmid vector derived from Corynebacterium glutamicum or Brevibacterium lactofermentum. The method is used for transforming a high temperature-resistant coryneform microbe. The present sequence encodes a Brevibacterium lactofermentum p48K protein sequence, which is used in an example from the present invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,55e-16 Length: 4447
Score: 318.00 Matches: 100
Percent Similarity: 47.45% Conservative: 58
Best Local Similarity: 30.03% Mismatches: 134
Query Match: 15.86% Indels: 42
DB: 6 Gaps: 14
US-10-007-527A-2 (1-379) x ABL49734 (1-4447)
QY 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
DB 1396 ATGTACAAGATCAACCAATAGTAAGGCTCTGGCGGGGTGCATAGTG-CGCAGGACCGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProIleGlySerGlyPheGlyLeuAr 71
DB 1455 AGCTGTGTGGGTGCTCTGTCTGTCTAACCGT-----GCTTCGAGTTTGGGGTCTGCA 1508
QY 71 gSerCysGlyLysGlyTTPilleCysProCysCysAlaGlyLysValGlyAlaHisArgal 91
DB 1509 AAACCTCACTCTCGCTGGGGGTCTCTCTCTGCTGAATTTGGAAGTCTATGGCGAACGCGC 1568
QY 91 aspGluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
DB 1569 CATTGAGCTGGCTATTGCTACTAAGATCACTTGGCGCGGTGGCGGCTCATGATGTT 1628
QY 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130
DB 1629 TGTGGGCACTGTTCCGACACACCGCTCACAGTCACTTTGGCGAGGTTGAAGCGGGTATTA 1688
QY 130 rAlaAlaTrpIleSalalaThrAsnGlyArgArgTTPArgThrGluArgGluMetTyrGl 150
DB 1689 GACTGCGTACTCTTCGATGTGTGAACATCTCAGTGAAGAAAGACGTGCACGCTACGG 1748
QY 150 YCysAspGlyTyrValArgAlaValGluIleThrHisGly---LysAsnGlyTTPHisVa 169
DB 1749 GTGGAGACACCTATAGTACTATGAGTACAGACTCTTGGCGGAACGGTTGCGACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
DB 1809 GCACCGCAACATGCTGTCTTGGATCGTCCACTGCTGACGATGAACCTCAAGGCATT 1868
QY 189 eSerAspAlaMetPheAspArgTTPThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
DB 1869 TGAGGATTCATGTTTCCCGCTGGTCTGGTGTGAAGCCGGTATGGACGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgIleGlyGlyGluAlaAspGl 227
DB 1929 ACTGGTGAGCAGCGGGTCAAACTTGATCAGTGCTCTACCTGGGTGGAGCGCTGCCAA 1988
QY 227 nValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
DB 1989 A---ATGCCAACCTACCTCGCTAG-----GGCATGCTCAGGAACCTGACTGGCTC 2036
QY 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTTPGluIleAlaValAspAlaVa 267
DB 2037 CGTACTAAACCGCGTCTAAAGGGTCTGACACGCGGTTTCAGATG---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTTPArgGl 279
DB 2094 GCGCATCAACGCGCGCGGAGGAGATGATGACGCTGTTTGGTGGCTGGTGGCTGGA 2153
QY 279 uPheGluPheGlySerMetGlyArgAlaAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
DB 2154 GTATGAGTGTGCTTCTAAACACCTGCTGTCG---TCATTGTCACGTGGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu-- 315
DB 2205 GCGTGTGTTGGGCACTGATTACATAGACGCTGATGTGCTGTAATGGAAGAGACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
DB 2265 GTACAAGCTCGCGGTCTGGAGCACCAGCGGTCGAATCAACCGCGTGTGCTGTGC 2324
QY 323 alleleProAlaArgSerTTPMetMetIleArgThr 335
DB 2325 TTTGTGAAGCCCGATGATTGGAACCTGATTGCT 2361
RESULT 12
ABL49734
ID ABL49734 standard; DNA; 4447 BP.
XX
XX ABL49734;
AC
XX
XX 29-AUG-2003 (revised)
DT 29-MAY-2002 (first entry)
XX
XX Brevibacterium lactofermentum p48K nucleotide sequence SEQ ID NO:7.
XX
XX Kanamycin resistant gene; mutation; high temperature resistance;
XX coryneform microbe; transformation; gene; ds.
XX


```

1569 CATTGAGCTGGCTATTGCTACTAAGAATCACTTGGCGGGGTGGCGGCTCATGANGTT 1628
111 rMet-----ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1629 TGTGGCACTGTTTCGACACACCGCTCAGCTCATTTGGCGGCTGTAAGCGGGTATTAA 1688
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
130 rAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTyrGl 150
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1689 GACTGGCTACTCTCGATGTGTGAACATCTCAGTGGAGAAAGACGTGCACGGTACGG 1748
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
150 yCysAspGlyTyrValArgAlaValGluLeuThrHisGly---LysAsnGlyTrpHisVa 169
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1749 GTGGAGCACACCTATAGTGACTATGAGTGCACAGACTCTTGGCGGAACGGTGGCACTT 1808
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnLeuLeuGluSerPh 189
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1809 GCACCGCAACATGCTGTGTCTTGGATCGTCCACTGTCTGACATGACACTCAAGGCATT 1868
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1869 TGAGGATCCATGTTTCCCGCTGCTGCTGGTGTGTTAAGCGCGGTATGGACGGCC 1928
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGluAlaAspGl 227
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1929 ACTGGTGGACACCGGGTCAAACTTGATCAGGTCTTACCTGGGGTGGAGACGCTGGCAA 1988
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
227 nValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1989 A---ATGGCAACCTACCTCGCTAAG-----GGCATGCTCAGGAACCTGACTGGCTC 2036
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluLeuAlaValAspAlaVa 267
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2037 CGCTACTAAACCGCGCTTAAAGGTCGTACACGCGGTTTCAGATG---TTGGATATGTT 2093
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2094 GGCCGATCAAGCGACGCGCGCGGAGATATGACGCTGTTTGGTGGCTGGCGGTGA 2153
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
279 upheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2154 GTATGAGGTGGTGTCTTAAACCTCGGTTCG---TCTTGTTCAGCTGGG-----GCTAA 2204
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu-- 315
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2205 GCGTGTCTTGGGATTGATTACATACAGCTGATGTACGCTCGTGAATGGAAGAACT 2264
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
316 -----GluSerAlaProValMetValAl 323
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2265 GTACAAGCTCGCGGTCTGGAAGCACCGGAAACCGGTCGAATCAACCCGGTGTCTG 2324
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
323 alleleProAlaArgSerTrpMetMetIleArgThr 335
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2325 TTTGGTGAAGCCCGATGATTGGAAACTGATTCAGTCT 2361
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 14
ADB66212
ID ADB66212 standard; DNA; 4447 BP.
XX
AC ADB66212;
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE B. lactofermentum replication control region in plasmid pSPK6.
XX
KW L-arginine production; coryneform bacteria; lysE; arginine repressor;
KW argR; liver function promoting agent; amino acid infusion;
KW amino acid pharmaceutical; replication control region; plasmid pSPK6; ds.
XX
XX Synthetic.
OS Corynebacterium glutamicum.
XX
XX Key Location/Qualifiers
XX 1318..2601
FT

```

```

/*tag= a
/partial
/Note= "This sequence lacks a start codon"

```

US2003113899-A1.

19-JUN-2003.

17-JUL-2002; 2002US-00196232.

25-JUL-2001; 2001JP-00224586.

(AJIN) AJINOMOTO CO INC.

Yamaguchi M, Ito H, Gunji Y, Yasueda H;

WPI; 2003-708853/67.

P-PSDB; ADB66179.

A microorganism comprising enhanced expression of the lysE gene is useful for enhanced production of L-arginine.

Example 1; Page 17-19; 36pp; English.

The present invention relates to a method for producing L-arginine in a microorganism (e.g. coryneform bacteria) that has L-arginine producing ability and has been modified for enhanced expression of the lysE gene. The microorganism is also modified so that an arginine repressor (argR) does not function normally. The method of the invention is useful for the enhanced production of L-arginine which is useful in liver function promoting agents, amino acid infusion and comprehensive amino acid pharmaceuticals. The present sequence represents Brevibacterium lactofermentum replication control region in plasmid pSPK6.

Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,55e-16	Length:	4447
Score:	318.00	Matches:	100
Percent Similarity:	47.45%	Conservative:	58
Best Local Similarity:	30.03%	Mismatches:	134
Query Match:	15.86%	Indels:	42
DB:	10	Gaps:	14

US-10-007-527A-2 (1-379) x ADB66212 (1-4447)

QY	34	LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal	53
DB	1396	ATGTACAGATCACCAATAGTAAGGCTCTGGCGGGTGCATAGGTG-CGCAGGGACGA	1454
QY	54	Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr	71
DB	1455	AGCTGTGGCGGTGCTCTGTGTCGTCTAACCGT-----GCTTCGACGTTTGAGGCTTCGA	1508
QY	71	gSerCysGlyLysGlyTyrIleCysProCysAlaGlyLysValGlyAlaHisArgAl	91
DB	1509	AAACTCTACTCTCGCTGGGGTCACTCTCTGGCTGAATTGGAAGTCATGGCGGAACCGCG	1568
QY	91	aAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh	111
DB	1569	CATTGAGCTGGCTATTGCTACTAAGAAATCACTTGGCGGGTGGCGGCTCATGATGTT	1628
QY	111	rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe	130
DB	1629	TGTGGGCACCTGTTTCGACACACCGCTCACAGTCATTTGGCGGCTTGAAGCGGGTATTAA	1688
QY	130	rAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTyrGl	150
DB	1689	GACTGCTACTCTCTCGATGTGTGAACATCTCAGTGGAGAAAGAACGTGCACCGTACGG	1748
QY	150	yCysAspGlyTyrValArgAlaValGluLeuThrHisGly---LysAsnGlyTrpHisVa	169
DB	1749	GGTGGAGCACACCTATAGTGACTATGAGTGCACAGACTCTTGGCGGAACGGTGGCACTT	1808

Sat Oct 23 17:44:52 2004

```

QY 227 GlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySer 246
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6520 AAA---ATGCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAAGTCACTGGC 6473
QY 247 GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAla 266
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6472 TCCGCTACTAAACCCGGTCTAGGGGTCTACACGCCGTTTCAGATG---TTGGATATG 6416
QY 267 ValGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArg 278
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6415 TTGCCCGATCAAAGCGACGCCGCGAGGATATGACCGCTGTTTGGTGGCTCGGTGGCGT 6356
QY 279 GluPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAla 298
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6355 GAGTATGAGGTGGTCTTAAACCTCGGTTTCG---TCCTGGTCACGTGGG-----GCT 6305
QY 299 ArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu 315
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6304 AAGCGTGCTTTGGGCATTGATTACATAGACGCTGATGTACGTCGTGTAATGGAGAGAA 6245
QY 316 -----GluSerAlaProValMetVal 322
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6244 CTGTACAAGCTGCCGGTCTGGAAGCACCGGTCGATCAACCCGCGTTGCTGTT 6185
QY 323 AlaIleIleProAlaArgSerTrpMetMetIleArgThr 335
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
6184 GCTTTGGTGAAGCCCGATGATTGGAAACTGATTCACTCT 6146

```

Search completed: October 23, 2004, 04:15:20
Job time : 680 secs

result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	337	16.8		7560	4	US-09-754-112A-2	Sequence 2, Appli
C 2	337	16.8		12738	4	US-09-754-112A-1	Sequence 1, Appli
C 3	326	16.3		4447	3	US-09-521-668B-17	Sequence 17, Appl
C 4	318	15.9		4447	3	US-09-521-668B-19	Sequence 19, Appl
C 5	214	10.7		5648	3	US-09-371-008-1	Sequence 1, Appli
C 6	122.5	6.1		441529	3	US-09-103-840A-1	Sequence 1, Appl
C 7	122	6.1		1686	4	US-09-252-991A-155	Sequence 155, App
C 8	122	6.1		1803	4	US-09-252-991A-170	Sequence 170, App
C 9	122	6.1		2610	4	US-09-252-991A-175	Sequence 175, App
C 10	120.5	6.0		4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 11	119	5.9		1428	4	US-09-252-991A-8370	Sequence 8370, Ap
C 12	119	5.9		1910	3	US-09-367-051-2	Sequence 2, Appli


```

QY 228 ValLeuAlaAlaTyrLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCGAGTACATCGCAAGACCCAGAGCGGAGCGCGCCCTCGAAGCTGCC 793
QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle-----262
Db 792 CGGCGCGACCTCAAGACGCGGAGCGGCGGGAACGTCGCCCGTTCGAACCTCCTCGGACGG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla-----273
Db 732 ATCGGGGACCTGACCGCGGCATGACCGAGGACGACGCGCGCGGCTCGGCTCGCTGGAG 673
QY 274 -----LeuGluLeuTrpArgGluPheGlySerMetGlyArgArgAlaIleAla 291
Db 672 TGGAACTCTCGCGCTGCGACGAGTACGAGCGGCAACCGCGGAGCGCGGCGCCATCGAA 613
QY 292 TrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCGGCTACCTCGCGGAGATGCTCGGCTCGACGGCGGCGGACACCGAGGCC 559

RESULT 3
US-09-521-668B-17
; Sequence 17, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAUKHIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521,668B
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: JP 11-69896
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 4447
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1318)..(2598)
US-09-521-668B-17

Alignment Scores:
Pred. No.: 2,36e-23 Length: 4447
Score: 326.00 Matches: 101
Percent Similarity: 47.75% Conservative: 58
Best Local Similarity: 30.33% Mismatches: 133
Query Match: 16.26% Indels: 42
DB: 3 Gaps: 14

US-10-007-527A-2 (1-379) x US-09-521-668B-17 (1-4447)

QY 34 LeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
Db 1396 ATGTACAGATACCAATAGTAGGCTCTGCGGGGTGCCATAGGTG-GCGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr 71
Db 1455 AGCTGTTGCGGTGCTCGTGGTCTGCTACGGT-----GCTTCGAGTTGAGGCTCTGCA 1508
QY 71 gSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AAACCTCTCACTCTCGCTGGGGGTCAACCTCTGGCTGAATTGGAAGTCATGGCGNACGCG 1568
QY 91 aAspGluIleSerGlnValAlaAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATGAGCTGGCTATTGCTACTAAGAAATCATCTTGGCGGGGTGGCGCGCTCATGATGTT 1628

```

```

QY 111 rMet-----ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130
Db 1629 TGTGGGCACTGTTTCGACACAAACCGCTCACAGTCATTTGCGCAGGTTGAAGCGGTATAA 1688
QY 130 rAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTyrGl 150
Db 1689 GACTGGGTCTCTTCGATGGTGAACATCTCAGTGGAGAAAGAAAGACGTGCACGGTACGG 1748
QY 150 yCysAspGlyTyrValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa 169
Db 1749 GGTGGGACACACCTATAGTACTATAGGTTCACAGACTCTTGGCGCAACGGTTGGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 1809 GCACCGCAACATGCTGTTCTTCTTGGATCGTCCTGATCGTCGACGATGAACCTCAAGCGGT 1868
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCATGTTTCCGCTGCTGCTGCTGTTAAAGCGCGGTATGACGCGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGluAlaAspGl 227
Db 1929 ACTGCTGAGCAGCGGGTCAAACTTGATCAGTGTCTACTGGGTGGAGCGCTCGGAA 1988
QY 227 nValLeuAlaIleTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCACTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGGCTC 2036
QY 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
Db 2037 CGCTACTAAACCGCGTCTAAAGGGTCTGACACGCGGTTTCAGATG---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
Db 2094 GCGCATCAAGCGAGCGCGCGAGGATATGACGCTGTTTGGTGGCTCGGTGCGTGA 2153
QY 279 uPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
Db 2154 GTATGAGGTTGGTCTTAAACACCTCGCTTCG---TCCTGGTCACGTGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGluIle-----ValGluGlnGlu--- 315
Db 2205 GCGTCTTTGGGCATTGATTACATAGACGCTGATCGTCGTAAGTGAAGAAAGAACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGCTCTGGAAGCACCGGACCGGTGAATCAACCGCGTGTCTGTTC 2324
QY 323 alleProAlaArgSerTrpMetMetIleArgThr 335
Db 2325 TTTGGTGAAGCCCGATGATTGGAACCTGATTCACTCT 2361

RESULT 4
US-09-521-668B-19
; Sequence 19, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAUKHIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521,668B
; CURRENT FILING DATE: 2000-03-08
; PRIOR FILING DATE: JP 11-69896
; PRIOR APPLICATION NUMBER: JP 11-69896
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 4447
; TYPE: DNA

```

```

; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1318)..(2598)
US-09-521-668B-19
Alignment Scores:
Pred. No.: 1.59e-22 Length: 4447
Score: 318.00 Matches: 100
Percent Similarity: 47.45% Conservative: 58
Best Local Similarity: 30.03% Mismatches: 134
Query Match: 15.86% Indels: 42
DB: 3 Gaps: 14

US-10-007-527A-2 (1-379) x US-09-521-668B-19 (1-4447)
Qy 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
Db 1396 ATGTACAGATACCAATAGTAAGGCTCTGGCGGGTGCATAGTG-GCGCAGGAGCA 1454
Qy 54 Asn-----GlyValThrIleVal-AsnGlyProIleSerGlyPheGlyLeuArg 71
Db 1455 AGCTGTTGGGTTCTCTGCTCTAAGCT-----GCTTCGAGTTTGAGGGTCTGCA 1508
Qy 71 gSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AACTCTCACTCTCGTGGGGTCTCTCTGCTGAATTGGAAGTCATGGCGAAGCGCG 1568
Qy 91 aAspGluIleSerGlnValAlaAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATGAGCTGGCTATTGCTACTAAGATCACTTGGCGGGTGGCGGCTCATGATGTT 1628
Qy 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130
Db 1629 TGTGGGCACTGTTCGACCAACCGCTCACAGTCATTGCGCAGGTTGAAGCGGTATTA 1688
Qy 130 rAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgTrpGluArgGluMetTrpGl 150
Db 1689 GACTGCTACTCTTCGATGGTGAACATCTCAGTCGGAAGAAAGAACGTCGACGGTAC 1748
Qy 150 yCysAspGlyTyrValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa 169
Db 1749 GTGGGAGCACCTATAGTACATGATGAGTCACAGACTCTTGGCGGACGGTTGGCATT 1808
Qy 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 1809 GCACCGCAACATCTGTGTTCTTGGATCGTCCACTGTCTGACGATGAATCAAGGCATT 1868
Qy 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCATGTTTCCCGCTGCTGCTGTTGTTAAGCGCGGTATGACGCGCC 1928
Qy 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGlyGluAlaAspGl 227
Db 1929 ACTGCGTGACACGGGTCAAACTTGATCAGGTGCTACCTGGGTGGAGACGCTGGAA 1988
Qy 227 nValLeuAlaAlaTrpThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGAGTGC 2036
Qy 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
Db 2037 CGCTACTAATAACCGCGTCTAAAGGGTCTGACACCGCGTTTCAGATG---TTGGTATG 2093
Qy 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
Db 2094 GGCCGATCAAGCAGCGCGCGGAGGATATGACGCTGTTTGTGGTCTCGGTGGCGTGA 2153
Qy 279 uPheGluPheGlySerMetGlyArgAlaAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
Db 2154 GTATGAGGTTGGTCTTAAACACCTCGCTTCG---TCTTGTGTCACGTGGG-----GCTAA 2204
Qy 299 sAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu-- 315

```

```

Db 2205 GCGTGTCTTGGCGATTGATTACATAGACGCTGATGTACGTCGTAATGGAAGAACT 2264
Qy 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGGTCTGGAAGCACCGGACGGGTGATCAACCCCGTGTGTGTGC 2324
Qy 323 alleleProAlaAArgSerTrpMetMetIleArgThr 335
Db 2325 TTTGTGAAGCCCGATGATTGAAACTGATTCACTCT 2361
RESULT 5
US-09-371-008-1/c
; Sequence 1, Application US/09371008
; Patent No. 6127174
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: SUGIYAMA, Masakazu
; APPLICANT: YOKOZAKI, Kenzo
; TITLE OF INVENTION: Plasmid Derived from Gluconobacter Bacteria
; TITLE OF INVENTION: and a Vector
; FILE REFERENCE: OP873
; CURRENT APPLICATION NUMBER: US/09/371,008
; CURRENT FILING DATE: 1999-08-10
; EARLIER APPLICATION NUMBER: JP 10-227227437
; EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5648
; TYPE: DNA
; ORGANISM: Gluconobacter oxydans
US-09-371-008-1
Alignment Scores:
Pred. No.: 1.25e-11 Length: 5648
Score: 214.00 Matches: 91
Percent Similarity: 40.00% Conservative: 45
Best Local Similarity: 26.76% Mismatches: 158
Query Match: 10.67% Indels: 46
DB: 3 Gaps: 16

US-10-007-527A-2 (1-379) x US-09-371-008-1 (1-5648)
Qy 15 ProProValLeuValSerSerAspLysArgGlyIleArgHisGluLeuArgProLysLeu 34
Db 3561 CCCCCAGAGCAAAACAGCAGCAAAATGGCGGTATCCAGGCTCTTAGAAGGCTTGT 3502
Qy 35 GlnGlnIleThrThrSerGluThrPheAsnAlaCysGly----- 47
Db 3501 CACGATCATGACGTGAGATGTTCCGATCCGCGCTTGTGGTTGCTCGCAAGATACTGCC 3442
Qy 48 -----ArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLys 63
Db 3441 CACATTTCTCCGCGCTTGTGCTGCTGGGATGGATCAACGCGTTG----- 3394
Qy 64 GlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysProCysAla 83
Db 3393 -----GCCGTGATGGGTCAGCATTCGCGCTCTGCT--CGTTGCCCTCACTGTGCG 3343
Qy 84 GlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal----- 98
Db 3342 -----CCTACCGGCATCTGATGTGCTGCCCGGTTGGTGGCGTCTCTCAAA 3295
Qy 99 ---AlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThrArgHisThr 117
Db 3294 GCGGTTTCATCCATGATGATGACGGCT---GCATTCGCCACATGCACTATGCGCATGAC 3238
Qy 118 AlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThr 137
Db 3237 CGAAACACTCGCTGTCAACATCGCGACGGTTCAGACTAAGGCTATTTCGCGATGCGAG 3178
Qy 138 AsnGlyArgArgTrp---ArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg 156

```

```

Db      3177  CGTGGTGGCTTGTGGAAACCGGCTGCTCAGGCAC-----GGTCTGGCTCGCTTATCCGG 3124
QY      157  AlavalgluilethrHisGly---LysAsnGlyTyrHisValHisValHisAlaLeuLeu 175
Db      3123  GTCCTTGGTGGTGGCTGGGCGGTAGACGGGCTGGCATCTGCATGTCCATGGATTGTT 3064
QY      176  MetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAsp 195
Db      3063  AUTCACCGGACGCGCAAGGCGCGT-----AGTCCGGCGCAGTCTCAGAGT 3010
QY      196  ArgTyrThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGly 215
Db      3009  CGCTGGATCAGCATGTGTGGCAAGGCTGTCATATCGCGTT-----GGA 2965
QY      216  LeuAspValArgLysIleGlyGlyGluAlaAspGlnValLeuAlaIleLeuThr 234
Db      2964  TCAGGGCAGAGGTGTTCCTGGATATCTGAAGACAAAGGCTTCTGATTTATGGGTGCGCA 2905
QY      235  LysIleAlaSer---GlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArg 253
Db      2904  GATCTTCGAGTGTGGGAATGCTTCTGAGATGGCGCGGATGGAAGAAACCGGGAAG 2845
QY      254  HisGlyAsnArg-----AlaProTrpGluIleAlaValAspAlaValGlyGlyAspPro 271
Db      2844  CGCCCTAATCGCCTGAATGTTCTCT---GAGTTGCTTGGCTTGGCGCTGAAGGTGACCGA 2788
QY      272  GlnAlaLeuGluLeuTyrArgGluPheGluPheGlySerMetGlyArgAlaIleAla 291
Db      2787  CTTGCTGCCAGAGTACCGCAAGCATGTTCTTGCCTGTGTCAGGTCAACGCTCTGCTGTG 2728
QY      292  TrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGlu---LeuThrAspAlaGln 310
Db      2727  GTTGCTCCCGATTTGAAGGGATCTTAGATCTGGACTTCGAGACATTCGACAGACAGCAA 2668
QY      311  IleValGluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSerTyr 330
Db      2667  GAGCGGAGCTTGAAGCGCAGCAGCAAAAGCCTGTGTGGCTTCTTCGGTGTGCTTGG 2608

```

RESULT 6

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

```

Alignment Scores:
Pred. No.: 614 Length: 4411529
Score: 122.50 Matches: 96
Percent Similarity: 38.19% Conservative: 35
Best Local Similarity: 27.99% Mismatches: 138
Query Match: 6.11% Indels: 76
DB: 3 Gaps: 18

```

US-10-007-527A-2 (1-379) x US-09-103-840A-1 (1-4411529)

QY 46 CysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65

```

Db      224503  TGCACCGACCAAGCGCGCTGTGCACGAAGTTGCGCTGTTGTCGATCCGACGGGCTC 224444
QY      66  GlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysProCysCysAlaGlyLys 85
Db      224443  GAGACGGTCACTTGTAGCTTGGCCCGAGTTGACCGGAGACCCCGTA-----CGT 224393
QY      86  ValGlyAlaHisArgAlaAspGluIleSerGlnValAlaHisGlnLeuGlyThr--- 104
Db      224392  CGAGGAGTCCAT-----GACCACCGTGGCGAAGTCTCAAGCACGCTAGTAGTGCAATT 224339
QY      105  ---GlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHis 123
Db      224338  CGGGGTCTGTGGCGCGCGACACCATGCCATGCCATGCGGGAAGGT---CTTTTGCAC 224282
QY      124  AspLeuTyrThrGly---LeuSerAla---AlaTyrLysAlaAlaThrAsnGlyArgArgT 142
Db      224281  GACGCGGTGCGGGTGTGTCATCTGATCGCTGCGCTCGT-----CGCGGT 224234
QY      142  TrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg---AlaValGluIleThr 161
Db      224233  GSCCACCAGCAATCACCGGTCCGAGCTCGACTTGTGTCGCGCGCATCGATGTTTAC 224174
QY      162  HisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal 181
Db      224173  CGCGTGTGTGGCGGCC-AGAACTCAACCGCGCGCACGCTACCTGAGTCGGGCTCATCC 224115
QY      182  SerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyr----- 197
Db      224114  AGTAG-AACTTG-----CCCGCACCGCGCGGATGGAAGATGTGATGGCGCTTCG 224062
QY      198  -----ThrSerLysLeuValSerLeuGlyPheAlaAlaPro 209
Db      224061  GCCATCGCCCGTCCGGGATCGCACCGAGTTGATGATCGGTCGGTGTGCCACACCG 224002
QY      210  ---LeuArgAsnSerGlyLeuAspVal-----ArgLysIleGlyGlyAlaAsp 226
Db      224001  ACCATCGCGCGTCCGACCGGATTGAAGCTCGCGCGCGGAAGCGCGCGGCCCAACAC 223942
QY      227  GlnVal-----LeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGly 241
Db      223941  GATGTCGATCCTCCCGGCAACCGGGCCCGCTCTACCGGCACTGCCGCGGTCAGGC 223882
QY      242  MetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrGlu 261
Db      223881  GGGGCGCGCGAGCGGTGAGTGGCGGCGCGCGCCCAACGGGTCAAGTCCCGCGGC--- 223825
QY      262  IleAlaValAspAlaValGlyGlyAsp---ProGlnAlaLeuGluLeuTyrArgGluPheG 281
Db      223824  ---AGCGCCCAACGGGTGCGGTGCGCGCGCAACA----- 223794
QY      281  uPheGlySerMet-----GlyArgArgAlaIleAlaTyrSerArg----- 294
Db      223793  ---GTCACCTCGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 223738
QY      295  -----GlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGln-- 310
Db      223737  CGGGATCGCGCGGAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 223678
QY      311  -----IleValGluGlnGluSerAlaProValMetValAlaIleIle 325
Db      223677  CGCAACCACTGCGGTATGTTCAAGCGCGCAATAGACTTCCCTCGTCCGACCGCTGCT 223618
QY      325  eProAlaArgSerTyrMetIleArgThrCysAlaProTyrValPheGlyGluIleLe 345
Db      223617  TCAGCGCGA-----CATAAATGCTTTTCCC-----TCGGGTGACGTCT 223579
QY      345  u 345
Db      223578  T 223578

```

RESULT 7

US-09-252-991A-155

; Sequence 155, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 155
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-155

Alignment Scores:
Pred. No.: 0.00686 Length: 1686
Score: 122.00 Matches: 98
Percent Similarity: 33.49% Conservative: 47
Best Local Similarity: 22.63% Mismatches: 166
Query Match: 6.08% Indels: 124
DB: 4 Gaps: 19

US-10-007-527A-2 (1-379) x US-09-252-991A-155 (1-1686)

QY 6 AlaGluHisLeuSerGlyAspArgProValLeuValSerSerAspLysArgGly 25
DB 474 GCTGACGACAACTGTCGGCTCGCCA-----GAAACGCGAGGA 515
QY 26 IleArgHis-----GluLeuArg---Pro 32
DB 516 CTTGACGACAACTGTCGGCTCGGCGCGCTGTTGAAGAGCGCGTGTGGAGAA 575
QY 33 LysLeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProLysSerGly 52
DB 576 CGGCTTCGGCGCGCTCTCCAGCAGACCGAAGCTGTCGGCTGGCGCGGCTGTGTG 635
QY 53 ValAsnGlyValThrIleValLeuGlnGlyProLysGlySerGlyPheGlyLeuArgSer 72
DB 636 GTTGACACAGCAAAATTTGTTGAAGCGCCCGG-----GCTGCTCTTCGTAC 683
QY 73 CysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
DB 684 TCCGGCGCAACCTTCAGCAGACACCGGGCAACTGTT-----CACCGGCT--- 728
QY 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMet 112
DB 729 -----GTTCTGCTGCGCACTGTCGGCGGAGCGCAGCAGCGCTGCTCGAACG 776
QY 113 ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrPheGlyLeuSerAlaAla 132
DB 777 CTTGCGGACCTGTTCTGGAAG-----GCG 803
QY 133 TrpLysAlaAla-----ThrAsnGlyArgGlyTrpArg 143
DB 804 TGGCGCGAGCTGGTGGCGCCACAGCAGCGGCTACTCGCGCGCTCGGTGGCGGA 863
QY 144 ThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluIleThrHisG1 163
DB 864 -----GTTGCGCGGAGTGTGCGCGGAGCGGAGCGGCTCTCGGCCAA 905
QY 163 YLysAsnGlyTrpHisValHis---ValHisAlaLeuLeuMet-----PheSe 178
DB 906 CCGCGATGCGCTGATCGCCCGCGCGAGCGGCTGGCGAGCTGGCGAGGTTTCGC 965
QY 178 rGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpTh 198
DB 966 CGGCGAGATCGGTTACGCTTGGCGGACAGCGAGCGGCGCCCGCTCTCGGCGAAGAGCCT 1025

QY 198 rSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspVa 218
DB 1026 GAGCAAGCTCTGCTCCCTCGGC-----GCGCGAGCGCGCGAGCTGGAGTT 1073
QY 218 lArGlyIleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSe 238
DB 1074 CAGCGCGAGCGCGGATCGCGAGGAGCGGCTGCCAGCCTTGTGCGCGGTCGCGGA 1133
QY 238 rGlyValGlyMetGluValGlySer-----GlyAs 248
DB 1134 GGGCTCGCGAGGAGGTGGAAGCGTGGCGGAGAGCGCTTCCCGATGCGTGGCGGA 1193
QY 248 pGlyLysSerGly-----ArgHisGlyAsnArgAlaProTrpGluI1 262
DB 1194 GCGGAAGAGATCGCGCGCGCGCTGCGCGCGGAGCGCTGCGAGGATCGC 1253
QY 262 eAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPh 282
DB 1254 CGGCTCGCGCGGATCGCGAGTGTGCGCGCATGTCCAGGTGGCGAGCGCTTCGAGTT 1313
QY 282 eGlySerMetGlyArgArg---AlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaG1 301
DB 1314 CCAGCGCGTGGCGAGTGTGCGCGCATGAAACGCGAGCGCTGTGCGGCGCAGCGGCG 305
QY 301 YLeuGlyAlaGlu-----GlyAs 305
DB 1374 GGTGACGAGGATGCTGCGCGCTGTCGAAACGCGAGCGGTGAAGGCGATCCGCGAGAT 1433
QY 306 -----LeuThrAspAlaGlnIleValGluGlnGluSerAl 318
DB 1434 TTTCTGTCACCCAGCGGAGATGCTGAGCATCGGAACTGCGCGAGAGGT-GCAACTGC 1492
QY 318 aProValMetValAlaIleIleProAlaArgSer-----TrpMetMetIleArg 334
DB 1493 GCCTGA-ATCGCGCGGAAAGCGCGGCTGAGTGTGGTGTGGAAGACACAGCGCG 1551
QY 334 g-----ThrCysAlaProTyrValPheGlyGluLeuGlu 346
DB 1552 GCGCAACAGGAGCGCTGACGACGCGTGTCTGCGCG-----AGCGCGCGCGACCTC 1605
QY 346 YLeuValGluAlaGlyAlaThrTrpGluAsnLeuArg 358
DB 1606 CCGGACCTCGCGCGGCTGTGTCGCGGCTGTGCGGCTGTGCGG 1642

RESULT 8
US-09-252-991A-170/c
; Sequence 170, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 170
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-170

Alignment Scores:
Pred. No.: 0.00756 Length: 1803
Score: 122.00 Matches: 98
Percent Similarity: 33.49% Conservative: 47
Best Local Similarity: 22.63% Mismatches: 166
Query Match: 6.08% Indels: 124

Db 2193 -----GTTCTGCTGGCCAGCTGGCGGAAGCAGCAGCGGCTGCTCGAACG 2146
Qy 113 ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAlaAla 132
Db 2145 CTTGTGGACCTGTTGCTGGAGG-----GCG 2119
Qy 133 TrpLysAlaAla-----ThrAsnGlyArgArgTrpArg 143
Db 2118 TGGCGCGAGCTGGTGGCGCCACACAGCAGCGCAGCTACTCGCGCGCTCGGTGGCGA 2059
Qy 144 ThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluLeuThrHisG1 163
Db 2058 -----GTCGCCCGCGAGCTGGCCCGAGCGCGCGAGCGGTCCTGGCCAA 2017
Qy 163 yLysAsnGlyTrpHisValHis---ValHisAlaLeuLeuMet-----PheSe 178
Db 2016 CCGCATGCGCTGCATCCCGCGCGCGCGCGCTGGCGAGCTGGCCAGGGTTGCG 1957
Qy 178 rGlyAspValSerGluAsnLeuLeuGluSerPheSerAspAlaMetPheAspArgTrpTh 198
Db 1956 CGGCAGAGATCGGGTACGCTGGCGCAGCAGCGCGCGCCCTCTCGCGGAGAGCCT 1897
Qy 198 rSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspVa 218
Db 1896 GAGCAAGCTGCTGGCCCTCGGC-----CGCGACCGCGCCAGACGCTGGAGTT 1849
Qy 218 lArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSe 238
Db 1848 CAGCGCCGAGCGCGCATGCGCGAGCAGCGCTGCCAGCTTGTCTGGCGCGTGGCGA 1789
Qy 238 rGlyValGlyMetGluValGlySer-----GlyAs 248
Db 1788 GGGGCTCGCGAGGAGTCCAAAGCGCTGGCGAGGAGCGCTGCCCGATGCGCTGGCGA 1729
Qy 248 pGlyLysSerGly-----ArgHisGlyAsnArgAlaProTrpLul1 262
Db 1728 GCGGAAGAGGATGCCCGCGCGCGCTGGCGCGCGGAGCGCTGCAGCGCATCGC 1669
Qy 262 eAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPh 282
Db 1668 CGGCTCGCGCGCATCGCAGTGGTCCGGCGCATGTCAGTGGCGCAGCGCTTCGAGTT 1609
Qy 282 eGlySerMetGlyArgArg---AlaIleAlaTrpSerArgGlyLeuArgAlaAlaG1 301
Db 1608 CCAGCGCGCTGGCGAGTCCCGGCCCATGAACGCGAGCGCTGCTGCGGCGCAAGCGGC 1549
Qy 301 yLeuGlyAlaGlu-----305
Db 1548 GGTGGAGCAGGAGATGCTCGGCTGTGTCGAACGACGACGCGTGAAGCGGATCCGCGAGAT 1489
Qy 306 -----LeuThrAspAlaGlnIleValGlnGlnGluSerAl 318
Db 1488 TTTCGTACCCACCGGAGATGCTCGACGATCCGGAACCTGGCGAGCAGGT-GCAACTGC 1430
Qy 318 aProValMetValAlaIlelleProAlaArgSer-----TrpMetMetIleAr 334
Db 1429 GCCTGA-ATCGCGCGGAAAGCGCGAGCGGCTGGAGTCGGGTGGTGAAGACAGCGCG 1371
Qy 334 q-----ThrCysAlaProTyrValPheGlyGluIleLeuG1 346
Db 1370 GCGCAACAGAGGCGGTGCAGGCGCTGCTGCGCG-----AGCGCGCGCGGACCTC 1317
Qy 346 yLeuValGluAlaGlyAlaThrTrpGluAsnLeuArg 358
Db 1316 CGCGACCTCGCGCGGGTGTGTCGCCAGCTCTGCGG 1280

RESULT 10
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Alignment Scores:
Pred. No.: 985 Length: 4403765
Score: 120.50 Matches: 96
Percent Similarity: 38.19% Conservative: 35
Best Local Similarity: 27.99% Mismatches: 138
Query Match: 6.01% Indels: 76
DB: 3 Gaps: 18
US-10-007-527A-2 (1-379) x US-09-103-840A-2 (1-4403765)
Qy 46 CysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65
Db 224612 TCCACCCAGCACCGCGGCTGTGCAGAACTTGTGTGTCATCCGACGCGCGTC 224553
Qy 66 GlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysProCysCysAlaGlyLys 85
Db 224552 GGAGACGGTCAACTGTAGCTTGGCGCGAGTTGACCGGAGCCCGTA-----CGT 224502
Qy 86 ValGlyAlaHisArgAlaAspGluIleSerGlnValAlaHisGlnLeuGlyThr--- 104
Db 224501 CGAGGAGTCCAT-----GACCACCGTGGCGAACTTCTCAAGCACCTAGTAGTGCATT 224448
Qy 105 ---GlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHis 123
Db 224447 CGGGGTCTGTGGCGCGCACCATGCCATCGACATCGAGAGGT---CTTTGTCAC 224391
Qy 124 AspLeuTrpThrGly---LeuSerAla---AlaTrpLysAlaAlaThrAsnGlyArgArgT 142
Db 224390 GACGCGTTCGGGTGATTTGTCACTGATGCGTGGGTGCT-----CGCGGT 224343
Qy 142 TrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluIleThr 161
Db 224342 GGCACCGAGCATCCCGTCCGAAAGCTCGACTTGTGTCGCGCGCATCGATGTTTAC 224283
Qy 162 HisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal 181
Db 224282 CGCGGTGTGGCGGCGC-AGAACTCAACCGCGCGCAGCTACTGAGTGGGCTCATCC 224224
Qy 182 SerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrp----- 197
Db 224223 AGTAG-AACTTG-----CCGCGACCGGCGGATGACGAATGTGATGGCGCTTCG 224171
Qy 198 -----ThrSerLysLeuValSerLeuGlyPheAlaAlaPro 209
Db 224170 GGCATCGCCGCTCGGCGATCGGCACCGAGTGTGATGACGATCGCTTGGCCACACCG 224111
Qy 210 ---LeuArgAsnSerGlyGlyLeuAspVal-----ArgLysIleGlyGlyGluAlaAsp 226
Db 224110 ACCATCGCGCGTTCGACCGGATTTGAACGTGGCGGCGCGAACCGCGCGGCGCCACAAAC 224051
Qy 227 GlnVal-----LeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGly 241
Db 224050 GGTGTGGATCTCCCGCGCAACGGGGCGCGCTCTACTCCGCACTGCGCGGCTCAGGC 223991

QY	242	MetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTirpGlu	261
Db	223990	GGGGGCGGCGCGAGCGGGTCAGGTGGCGGCGGCGCAACGGGTCAAGTGCCGGCGGC	223934
QY	262	IleAlaValAspAlaValGlyGlyAsp-ProGlnAlaLeuGluLeuTrpArgGluPheGlu	281
Db	223933	---ACGCCAACGGGTGGGTGGCGGCCCAACA---	223903
QY	281	uPheGlySerMet-----GlyArgArgAlaIleAlaTrpSerArg-----	294
Db	223902	---GGTCAACTGCGCGGCGGTTCAGCGGGTCCCGCTGTGCCCCCGGGGCCCGCGGACCTC	223847
QY	295	-----GlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGln--	310
Db	223846	CGGGATCGCGGGAGCCGCGTCCGGGTCCGCACCGCAACGCCAGCACCCAGCATCAGCA	223787
QY	311	-----IleValGluGlnGluGluSerAlaProValMetValAlaIleI	325
Db	223786	CGCAACCACTGGGCTATGTTCAAGCGGCGCAATAGACTCCCTCGTCAGCCCTGCT	223727
QY	325	eProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrValPheGlyGluIle	345
Db	223726	TCCAGCGCGA-----CATATTGCTTTCCC-----TCCGGTGACGTTCT	223688
QY	345	u 345	
Db	223687	T 223687	
RESULT 11			
US-09-252-991A-8370/c			
; Sequence 8370, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252.991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 8370			
; LENGTH: 1428			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-8370			
Alignment Scores:			
Pred. No.:			
Score: 0.011 Length: 1428			
Percent Similarity: 119.00 Matches: 80			
Best Local Similarity: 35.06% Conservative: 35			
Query Match: 24.39% Mismatches: 94			
DB: 5.94% Indels: 120			
Gaps: 16			
US-10-007-527A-2 (1-379) x US-09-252-991A-8370 (1-1428)			
QY	69	GlyLeuArgSerCysGlyLysGlyTrpIleCysProCysAlaGlyLysValGlyAla	88
Db	1361	GGCATCAGAAAGCTGTAGCTCAGCGC-----GGCCGAACAGAGTC	1323
QY	89	HisArgAlaAspGlu-IleSerGlnValValAlaHisGlnLeuGlyThrGlySerVal	108
Db	1322	GCCCGTAGGATCGGTGTAGCGCGGTCTAGCCGCTCTGGAAGGTAGGTCTGCTTGC	1263
QY	108	a-----MetValThrMetThrMetArgHisThrAlaGlyGln-----	120
Db	1262	TGAACGGCGTTTCGGGTTCGAACAGGTCTTCACCCCGAGGTTCAGCCCGCGGTGC	1203
QY	121	-----ArgLeuHisAspLeuTr	126

1202 TGAGCGGTAGTGTGCCGCCAGGTCCACAGAGTTCCAGAACCCGACCTTGCCGTGGC-TG 1144
126 pThrGlyLeuSer-:::-----AlaAlaTyrLysAlaAlaLath 137
1143 GCGCGGTCTGAGTGGTGTAGCCCGTGGTGTAGCGGTGTCTCAGGGTGGCTGCACGCG 1084
137 rAsnGlyArgArgTTPArgThrGluuArgGluMetTyr-Gly-:::----- 150
1083 TCGCGGCTCCAGGTGGCGTTCGGACAGTGTCTCCAGCGGGGATCATCGCCAGCAGCCCTGG 1024
151 -----CysAspGlyTyrValArg-----Ala-ValGluIleThrHisGlyL 164
1023 AAGTCGCGACAGTGTGTCCAGTACTGGCGCGCATCTGCTCTGGAAGTCGTAGCGGA- 965
164 ysAsnGlyTTPHisValHisValHisAlaLeuLeu-----MetPheS 178
964 -----CACGTAGGTGCCCTTGCAGGTCGAGCCCGAACTCCCGGTAGCGG 922
178 erGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpT 198
921 CTGGC-----CGGAAACGATAATCAGGCTCAGGTTCGACGCG----- 884
198 hrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAenSerGlyGlyLeuAspV 218
883 -----GCTGGTCTTTTACC 871
218 alArgLysIleGlyGlyGluAlaAspGlnValLeu-----AlaAlaTyrLeuThrLysL 236
870 TTGCGGAGGTGGCCAGTCCGGTGACGACGTGATCGATAGAGCCGTCGCGCTTTGCGCAG 811
236 leAlaSerGlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyA 256
810 ATGGTCCGCGTAGGCTCCGGGTCCGGCAA--GATCGCGCTTCGGGAAACTCGGCG 754
256 snArgAlaProTTPGluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluL 276
753 AT-----CTGGTTGGCGATCTGTATCCACCAGAAAG 724
276 euTTPArgGluPheGluPheGlySerMetGlyArgArgAlaIleAlaTTP----- 292
723 TCCAGGCGGAC-----GGNAAGTCCGCGAGCGGTGGTAGACAGGCCAGGGTA 673
293 -----SerArgGlyLeuArgAlaGlyLeuGlyAla----- 304
672 ACGTTCGCGCGGTTCGCGTCCGAGTCCGTATTCCGCCGCGTGGCGTTGTAGAACTGC 613
305 -----GluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProValM 321
612 TGGCGCAGTCGCGGTGGCGATCCGCTGGCTGGCTGGCTGGCCCGCGCACAGGCGC 553
321 etValAlaIlelleProAla 327
552 GGGTCGTCGTAGTTGGCGCT 533
RESULT 12
US-09-367-051-2
; Sequence 2, Application US/09367051
; Patent No. 6350577
; GENERAL INFORMATION:
; APPLICANT: OHTAKE, Hisao
; APPLICANT: KATO, Junichi
; APPLICANT: NAKAMURA, Yosuke
; TITLE OF INVENTION: PLASMIDS DERIVED FROM AMMONIA OXIDIZING BACTERIA AND
; TITLE OF INVENTION: USE OF THE SAME
; FILE REFERENCE: 20-4599P
; CURRENT APPLICATION NUMBER: US/09/367,051
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1910
; TYPE: DNA

Sat Oct 23 17:44:52 2004

```
; ORGANISM: Nitrosomonas sp.
; US-09-367-051-2
Alignment Scores:
Pred. No.: 0.0168 Length: 1910
Score: 119.00 Matches: 57
Percent Similarity: 42.42% Conservative: 41
Best Local Similarity: 24.42% Mismatches: 91
Query Match: 5.94% Indels: 42
DB: 3 Gaps: 12

US-10-007-527A-2 (1-379) x US-09-367-051-2 (1-1910)
QY 71 ArgSerCysGlyLysGlyTrpLeuPheCysProCysCysAlaGlyLysValGlyAlaHisArg 90
DB 763 CGCTTCTGCAAAAGCAACTTCGTGCGCTATGTGCGCTATTCGGCGGGTTCGAAATG 822
QY 91 AlaAspGluLeuSerGlnValAlaHisGlnLeuGlyThrGlySerValAla----- 108
DB 823 CTT-----ACATCGTATCTTCAACGGTATCAGATCGTGTACTGCTCAATCCGATCTG 876
QY 109 -----MetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeu 125
DB 877 AAGCCTTATCTGCTGACTTAAAGGTTAAGAT-----GGACCTGATTTCAGGAACGT 930
QY 126 TrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGlu 145
DB 931 TTTGTGCTATTCGCTAAAGCTATGCGCTCGATGTTCAAAATCTCGCAGAAATACTAATCAA 990
QY 146 ArgGluMetTyr-----GlyCysAspGlyTyrValArgAlaValGluLeuThr 161
DB 991 GCGCAGCGTTTCGTGAGTTTCTAAGTCTGAAGGTTGTTTCATTCGATTCGATGAGTAACT 1050
QY 162 HisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal 181
DB 1051 AACCGGGGCAACGGCTGGATCTCAGCGTCATATGATCTGTTGTGTAGGAAAAACCA 1110
QY 182 SerGluAsnLeuLeuSerPheSerAspAlaMetPheAspArgTrp----- 197
DB 1111 GATCAGTCGGCTTTTCGGGC-----GAATGGCGAGAGATTACC 1149
QY 198 ---ThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsn-----SerGly 214
DB 1150 GGTGATCTCAGCTGGTGTACGTC-----AGCCCTTAACTGATCCAGTTGACGGC 1200
QY 215 GlyLeuAspValArgLysIleGly-----GlyGluAlaAspGlnValLeu 229
DB 1201 TTTCTTGAGTTTGTAAAGTACGCTTAAAGTCTCTGATCTTGAGCTGGCCATCTTTTC 1260
QY 230 AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGlyAspGly 249
DB 1261 GAAGCGTAC-----AAGATTTTGTCTGTTCTCGCTGTC-----GATAGT 1302
QY 250 LysSerGlyArgHisGlyAsnArgAlaProTrpGluLeuAlaValAspAlaValGlyGly 269
DB 1303 CACGGATGTAGAGAGTTTCGGTTACCGGATGATCTTGGATGATGATGATGATGATGATG 1362
QY 270 AspProGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 280
DB 1363 TTGCCG---TTGCTTGAACCTTCTCTATCGATTT 1392

RESULT 13
US-09-252-991A-8473
; Sequence 8473, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8473
LENGTH: 3528
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8473

Alignment Scores:
Pred. No.: 0.0413 Length: 3528
Score: 119.00 Matches: 91
Percent Similarity: 34.65% Conservative: 41
Best Local Similarity: 23.98% Mismatches: 117
Query Match: 5.94% Indels: 134
DB: 4 Gaps: 18

US-10-007-527A-2 (1-379) x US-09-252-991A-8473 (1-3528)

QY 24 ArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGluThrPhe 43
DB 38 AGAGCGTAATAGTGTGATGCGACCGGAGGCCAACACCTGCCAGATCTGTGCGCCG 97
QY 44 Asn-----AlaCysGlyArgProIleSerGlyValAsnGly 55
DB 98 CGAAATCTGGCAGGCTACCGGATCTGTAGCTGACACAGCCGTCGACGGCTTGC GGCC 157
QY 56 ValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLys 75
DB 158 ATA-----CCGCCCGGC-----GCCCGCGCGGCATCAGAAAGCTGTAGCTGA 200
QY 76 GlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGlu-IleSe 95
DB 201 GGC-----GCCCAACAGGATGCGCCCGTAGGATCGGTGTAGC 239
QY 95 rGlnValValAlaHisGlnLeuGlyThrGlySerValAla-----MetValThrMe 112
DB 240 GCGGTGCTAGCGCTCTGGAAGTGTAGTCTGTGTTGCTGAACGCGGTTCGCGGTGA 299
QY 112 tThrMetArgHisThrAlaGlyGln----- 120
DB 300 ACAGGTCTTTCACCCCGAGGTCAGCCCGCGCGGTGGCTGAGCGGTAGCTGCGCGCCA 359
QY 121 -----ArgLeuHisAspLeuTrpThrGlyLeuSer----- 130
DB 360 GGTCCAGAGGTTCACAGAACCCAGCTTCCCGTGGC-TGGCGCGGTCTAGTGTGTAG 418
QY 131 -----AlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgTh 144
DB 419 CGGCTGCTAGCGGTGTGCTCAGGCTGGCTGCGAGCGGTCCAGGCTGCGCGGTG 478
QY 144 rGluArgGluMetTyrGly-----CysAspGly 153
DB 479 GCGACGTGCTTCCAGCGGCGGATCAGCGCGACCCCTGGAGTGGCGGAGCTGTTCAGG 538
QY 153 yTyrValArg-----Ala-ValGluIleThrHisGlyLysAsnGlyTrpHisValHisV 171
DB 539 TACTGGCGCGCGATCTGCTCTCGAAGTCTGAGGAGTCTAGCGGGA-----CACG 580
QY 171 alHisAlaLeuLeu-----MetPheSerGlyAspValSerGluAsnI 185
DB 581 TAGTGCCTTTCAGGTCGAGCCGCACTGCGCGGTAGCGGTGGC-----CGGGAACGA 634
QY 185 leLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuG 205
DB 635 TAATCGAGGCTCAGGTGCGAGCC----- 657
QY 205 lyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGluA 225
DB 658 -----GCTGCTCTTACCTTTCGCGAGGTTGCCAGTCCG 691
QY 225 laAspGlnValLeu-----AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetG 243

Db 692 GTGACGACGTGATCGATAGACGCGTTCGGCCCTTCGCGACGATCGCTCCGGCGTAGGCCTGC 751
QY 243 luValGlySerGlyAspGlyGlyValSerGlyValArgHi.sGlyAsnArgAlaProTrpGluIleA 263
Db 752 GGGTCGCGCAA---GATCGCGCTTCGGGNAATCGCGCAT--- 789
QY 263 laValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPheG 283
Db 790 -----CTGGTTGGCGATCTCATCCACAGAAAGTCCAGGCGGAC-----G 829
QY 283 lySerMetGlyArgArgAlaIleAlaTrp-----SerArgGlyLeuArgA 298
Db 830 GAAAGTCGCGCAGCGCTGGTAGACAGGCCCGAGGGTAACGTTGCGCGCGTTCCGGT 889
QY 298 laArgAlaGlyLeuGlyAla-----GluLeuThra 308
Db 890 CGCAGCTCGGTATTGCCCGCGCTGGTGTAGACCTGCTGGCGCAGTCGCGGTGGCG 949
QY 308 spAlaClnIleValGluGlnGluSerAlaProValMetValIleIleProAla 327
Db 950 ATGCCCGCTCGGCTCGGCTGGCGCGCGGCACAGGCGCGGTCGTCGTAGTTGGCGCT 1008

RESULT 14
US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: *Sorangium cellulosum*
US-09-335-409-1

Alignment Scores:		
Pred. No.:	5.13	Length: 68750
Score:	117.00	Matches: 95
Percent Similarity:	31.25%	Conservative: 45
Best Local Similarity:	21.21%	Mismatches: 161
Query Match:	5.84%	Indels: 149
DB:	3	Gaps: 18
US-10-007-527A-2 (1-379) x US-09-335-409-1 (1-68750)		
Qy	65	SerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysProCysCysala--- 83
Db	48236	TCGGAGGTTCCGGGCGCGACCTATGTGGCCAAAGGGTGCTTCTCCCGCATGTGCGCAG 48295
Qy	83	-----
Db	48296	CTTGGATGGCGGTCTTCGCCATTTCCTCGTAGGGCGATGAGCTGGACCCGCAACA 48355
Qy	84	-----GlyLysValGlyAlaHisArgAla----- 91
Db	48356	GCGGCTGTTGCTGGAGTGATGCTGGGAGCGATCGAGCGCTGGCCAGACCCCATGGC 48415
Qy	92	-----AspGluLeuSerGlnValValAlaHis-----GlnLeu 102
Db	48416	GCTGCGCGAGAGCGCCACGGGCGTGTTCTGGGCGATGTCGGAGCGAGCAGCGCGAGCG 48475
Qy	103	GlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeu 122

Db 48476 GGTCCAGGGCTCGACGACGACGCGGGTGTGTACGCCACCACCGCAACTGCTCAG 48535
Qy 123 HisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrp 142
Db 48536 CGTCGCGCTGGACGGGTGCCT-TCTTCTGGGTCTGCACGGCCGACGATGACGGTGG 48594
Qy 143 Arg-----ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 48595 ACACCGCTCTCTGCTGCTGCTGTGTGGCTGTGCACCTCGCTGCACAGCGCTGCGATTGG 48654
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 48655 GCAGTGCGACGAGCGCTGCGCGGG--TCCAGCGTGTCTTTGTCCGCGGTCAATC 48713
Qy 181 ValSerGluAsn--- 184
Db 48714 GTCGCGCGTCCGCGCATGCTTTGCTTTCGCCAGATGGCGGTGCAAGAGTTCCTCGGCC 48773
Qy 185 ----- 1leuGluSerPheSer 190
Db 48774 GTGCAAGCGGTTTTGCGGGCCGAGGCTGCGCGTGTGTGTCTCAAGCGGCTCCGT 48833
Qy 191 AspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAProLeu 210
Db 48834 GAGCGCAGCGCACCGC-----GACCCATCTCTGGCGGTGTGAGACGACGCGGATC 48887
Qy 211 ArgAsnSerGly-----GlyLeuAspValArgLysIleGlyGlyGluAlaAspGln 227
Db 48888 AACACAGATGGCCGACGACGCGGCTCACGGTGCCC-----AGCGGTCTCTGCCACGAC 48941
Qy 228 ValLeu----- 229
Db 48942 GGTGTTCTTAGCCGCGCTGCGCAAGCGGGGTGCGCGCGGAGGTGCTCAAGTTCGTG 49001
Qy 230 AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluVal-----GlySer 246
Db 49002 GAGTCCACGCGACGGGACAGCGGTGGGTGACCCGATCGAGTGCAGGCGTGGCGCG 49061
Qy 247 GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAla 266
Db 49062 GTGTACGGGCGG---GGCCGCCCGCGGAGCGCGCTCTGTGGTGGCGCTGTCAAGGCC 49118
Qy 266 ----- 266
Db 49119 AACCTCGGCCACTCGAGCGCGCGGGCTTTGGCGGGCTGTCTCAAGGTGCTCTTGGCG 49178
Qy 267 -----ValGlyAspProGlnAlaLeuGluLeu----- 276
Db 49179 CTGGAGCAGCAGCATTCGGCTCAAACGGAGCTCGACAGCTCAACCCGACATCCCG 49238
Qy 277 TrpArgGluPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeu 296
Db 49239 TGGCAGAGTCCACAGTG---GCCGTGTTCGCGAGGCGGTCCCCCTGGCGCGCGCGCG 49295
Qy 297 ArgAlaArg-----AlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleVal 312
Db 49296 CGCCGCGTCTGTGAGGCGTGAGCGCTTTCCGCGCTGAGCGGGACCAACGCGCATGTGGTG 49355
Qy 313 GluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSerTrpMetMet 332
Db 49356 TTGGAGGAGGCGCGCGGTGAGCGCTGTGGCGCGGCCCGCGAGCGCGGAGCTG 49415
Qy 333 IleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGluAlaGlyAla 352
Db 49416 TTCGTCTGTGCGCG-----AAGAGCGCGGGCGGTGATGCGCAGGCA 49460
Qy 353 ThrTrpGluAsnLeuArgAspHisLeu-----HistYrArgLeuProAlaAlaAspVal 370
Db 49461 GCC-----CGGCTGCGGGACCACTGGAGAAGCATGTCGAGTGTGGCTCGCGGATGTG 49514
Qy 371 ArgProProfileSerValArg 378
 ::: :::

49515 GCGTTCAGCCTGGCGACGACGCGC 49538

; Sequence 1, Application US/09568102

GENERAL INFORMATION:

APPLICANT: Ligon, James

APPLICANT: ZILKIE, ROSS
APPLICANT: CUR DEYON

APPLICANT: GOETZACH, ROBERT
TITLE OF INVENTION: GENES

FILE REFERENCE: 100-368611
CURRENT APPLICATION NUMBER: US/09/568,102

PRIOR APPLICATION NUMBER: 09/335,409

; NUMBER OF SEQ ID NOS: 30

; SEQ ID NO 1

; TYPE: DNA

US-09-568-102-1

Alignment Scores:

Score:	117.00	117.00
Machines:	31	258
Conservative:	45	45

BEST LOCAL SIMILARITY: 21.21%
 INDEXES: Match: 5 84%
 Indexes: 149

DB:

[illegible][illegible]

103EE

101

48415

-----Glen 102

[illegible]

103 GlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeu 122

48476 GGTGCAGGGCCCTCGACGACGACGGCGGCTTGCTGTACGGCACCCGCGCAACCTGCTCAG 48535

123 HisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrp 142

DB 48536 CGTCGCCGCTGGACGGCTGTCGT-TCTTCCTGGGTCTGCACGGCCCGACGATGACGGTGG 48594

QY
143 Arg-----ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160

Ddb
48595 ACACGGCCCTGCTCGTCTGGTGGGTGCACCCTGCCAGAGAGCC

[illegible]

D_B GCGAGTGCACGACCCTGGCCGCCGGC-TCCAGCGTCCTTTTGATCCCCCCCACCA...AATA-187.

Qy 181 valsergluash-----

[illegible]

QY	185	----	-----IleLeuGluSerPheSer	190
			::: :::	
Db	48774	GCTGCACAGCGCTTTGGCGGCGGAGCGCTGCGCGGTGGTCTCAAGCGCGTCCGT	48833	
QY	191	AspAlaMetPheAspA-GTTPTrSerIysLeuValSerLeuGlyPheAlaProLeu	210	
			::: :::	
Db	48834	GAGCGCAGCGCGACGCG------GACCCCATCTCTGGCGGTGTGTCAGGACACGCGATC	48887	
QY	211	ArgAsnSerGly-----GlyLeuAspValA-LgLyAlleGlyGlyGluAlaAspGln	227	
		:::		
Db	48888	AACACCATGATGCCGAGCAGCGGGCTACCGGTGCC------AGCGTCTCTCCACGACG	48941	
QY	228	ValLeu-----	229	
Db	48942	CGGTTGCTACGCCAGCGCGCTGGCGCAAGCGGCGGTGGCGCGCGAGTGCATTTCTGTG	49001	
QY	230	AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluVal-----GlySer	246	
		:::		
Db	49002	GAGTGCCACGGGACGGGACACGCGCTGGGTGACCCGATCGAGGTGCAGCGCTGGCGCG	49061	
QY	247	GlyAspGlyLysSerGlyValGHisGlyAsnArgAlaProTyrGluIleAlaValAspAla	266	
Db	49062	GTGTACGGGCGG------GGCGCGCCCGCGGAGCGCGCTCTGGCTGGCGCGCTGTCAAGGCC	49118	
QY	266	-----	266	
Db	49119	AACCTCGCCACCTGGAGGCGCGCGCGGCTTTGGCGGCGTGTCAAGGTGCTCTTGGCG	49178	
QY	267	-----ValGlyGlyAspProGlnAlaLeuGluLeu-----	276	
Db	49179	CTGGAGCAGCAGCATATCCGGCTCAACCGGAGCTCGACGAGTCAACCCGCATCCCG	49238	
QY	277	TyrArgGluPheGluPheGlySerMetGlyArgAlaIleAlaTyrSerArgGlyLeu	296	
Db	49239	TGGCAGAGCTGCCAGTG------GCCGTTGCCGACGGCGGTCCCTGGCCGCGCGCGCG	49295	
QY	297	ArgAlaArg-----AlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleVal	312	
		:::		
Db	49296	CGCCCGCTCGTCAGCGCTGAGCGCTTTTCGGCTGAGCGGACCAACCGGCATGTGGTG	49355	
QY	313	GluGlnGluGluSerAlaProValMetValAlaIleProAlaArgSerTyrMetMet	332	
Db	49356	TTGGAGGAGCGCGCGGTGGAGCTGTGGCGCGCGCCCGCGCGCGCGCGGAGCTG	49415	
QY	333	IleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGluAlaGlyAla	352	
Db	49416	TTGCTCCTGTGGCG------AAGAGCGCGCGCGCTGGATGCCAGSCA	49460	
QY	353	ThrTyrGluAsnLeuArgAspHisLeu-----HisTyrArgLeuProAlaAlaAspVal	370	
Db	49461	GCC-----CGGCTGGCGGACCACTGGAGAAGCATGTCAGGCTTGGCGCTCGCGGATGTG	49514	
QY	371	ArgProProIleIleSerValArg	378	
		:::		
Db	49515	GCCTTCAGCGCTGGCGACGCGCG	49538	

search completed: October 23, 2004. 07:50:26

Job time : 3743 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 05:46:24 ; Search time 607 Seconds
(without alignments)
3199.285 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGKDRPPVLVS.....HLHYLPADVRPPIISVRK 379

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10007527/runat_22102004_114118_26723/app.query.fasta_1.519
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -WAPR=US10007527 @CGN 1 1 544 @runat_22102004_114118_26723
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

* SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2005	100.0	1140	14	US-10-007-527A-1	Sequence 1, Appli
2	2005	100.0	1140	14	US-10-007-452-1	Sequence 1, Appli
3	2005	100.0	1140	17	US-10-415-562A-1	Sequence 1, Appli
4	2005	100.0	6334	14	US-10-007-527A-5	Sequence 5, Appli
5	2005	100.0	6334	14	US-10-007-452-5	Sequence 5, Appli
6	2005	100.0	6334	17	US-10-415-562A-5	Sequence 5, Appli
7	2005	100.0	9652	14	US-10-007-527A-7	Sequence 7, Appli
8	2005	100.0	9652	14	US-10-007-452-7	Sequence 7, Appli
9	2005	100.0	9652	17	US-10-415-562A-7	Sequence 7, Appli
10	2005	100.0	11241	14	US-10-007-527A-6	Sequence 6, Appli
11	2005	100.0	11241	14	US-10-007-452-6	Sequence 6, Appli
12	2005	100.0	11241	17	US-10-415-562A-6	Sequence 6, Appli
13	337	16.8	7560	9	US-09-754-112A-2	Sequence 2, Appli
14	337	16.8	12738	9	US-09-754-112A-1	Sequence 1, Appli
15	326	16.3	4447	9	US-09-835-381-5	Sequence 5, Appli
16	326	16.3	4447	15	US-10-196-232-5	Sequence 5, Appli
17	318	15.9	4447	9	US-09-835-381-7	Sequence 7, Appli
18	318	15.9	4447	15	US-10-196-232-6	Sequence 6, Appli
19	296	14.8	8500	16	US-10-296-947-7	Sequence 7, Appli
20	282	14.1	2401	9	US-09-826-191-6	Sequence 6, Appli
21	282	14.1	2401	14	US-10-263-666-6	Sequence 6, Appli
22	282	14.1	2401	14	US-10-261-481-6	Sequence 6, Appli
23	282	14.1	2401	14	US-10-261-942-6	Sequence 6, Appli
24	282	14.1	4005	9	US-09-826-191-2	Sequence 2, Appli
25	282	14.1	4005	14	US-10-263-666-2	Sequence 2, Appli
26	282	14.1	4005	14	US-10-261-481-2	Sequence 2, Appli
27	282	14.1	4005	14	US-10-261-942-2	Sequence 2, Appli
28	138	6.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
29	136	6.8	4368	17	US-10-437-963-33355	Sequence 33355, A
30	133	6.6	2256646	17	US-10-470-565-1	Sequence 96272, A
31	128.5	6.4	1763	17	US-10-437-963-96272	Sequence 1, Appli
32	125	6.2	2256646	17	US-10-470-565-1	Sequence 95384, A
33	123	6.1	1379	17	US-10-437-963-95384	Sequence 578, App
34	123	6.1	25582	17	US-10-322-281-578	Sequence 32408, A
35	122.5	6.1	1540	17	US-10-437-963-32408	Sequence 7827, Ap
36	122	6.1	2871	9	US-09-815-242-7827	Sequence 30310, A
37	122	6.1	2871	16	US-10-282-122A-30310	Sequence 25453, A
38	119	5.9	2427	16	US-10-282-122A-25453	Sequence 1, Appli
39	118.5	5.9	1386	17	US-10-437-963-48219	Sequence 1, Appli
40	118	5.9	135638	15	US-10-314-657-1	Sequence 7, Appli
41	117	5.8	68750	13	US-10-014-717-1	Sequence 12, Appli
42	116	5.8	1542	9	US-09-833-745-7	Sequence 6, Appli
43	116	5.8	15738	15	US-10-329-079-12	Sequence 53737, A
44	116	5.8	37360	15	US-10-329-079-6	
45	115.5	5.8	1661	17	US-10-437-963-53737	

ALIGNMENTS

RESULT 1
US-10-007-527A-1
; Sequence 1, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Rhodococcus AN12

Db	901	GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTGTGACGAGAGAAATCTGCCCGGTC	960
Qy	321	MetValAlaIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal	340
Db	961	ATGGTTGGCATCAATCCGGCGCGATCGTGAGTGAATTCGACCTTGCGGCTTACGTC	1020
Qy	341	PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis	360
Db	1021	TTGGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC	1080
Qy	361	LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys	379
Db	1081	TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCCCGCATATATATCGTTGCGAAG	1137
RESULT 2			
US-10-007-452-1			
; Sequence 1, Application US/10007452			
; Publication No. US20030093701A1			
; GENERAL INFORMATION:			
; APPLICANT: Tomb, Jean-Francois			
; APPLICANT: Bramucci, Michael G.			
; APPLICANT: Cheng, Qiong			
; APPLICANT: Kostichka, Kristy N.			
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors			
; FILE REFERENCE: C11709 US NA			
; CURRENT APPLICATION NUMBER: US/10/007,452			
; CURRENT FILING DATE: 2001-11-08			
; PRIOR APPLICATION NUMBER: 60/254,868			
; PRIOR FILING DATE: 2000-12-12			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 1			
; LENGTH: 1140			
; TYPE: DNA			
; ORGANISM: Rhodococcus AM12			
US-10-007-452-1			
Alignment Scores:			
Pred. No.:		8.33e-216	Length: 1140
Score:		2005.00	Matches: 379
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		14	Gaps: 0
US-10-007-527A-2 (1-379) x US-10-007-452-1 (1-1140)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	1	ATGACCGAGTAAAGTGTGAACACCTTTCGGCAAGACCGCGCTCCGCTCGTGTGCG	60
Qy	21	SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40
Db	61	TCCGATAGCCGCGCATCCGCGACGAACTGCGACCCAACTTCAACAAATCACCACGTCA	120
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	121	GAACATTTAACCTTGGCGCGCATTCGCGTGAACCGGTGTGACCATTTGTCAAC	180
Qy	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80
Db	181	GGTCCGAAAGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC	240
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100
Db	241	TGCTGTGCGGAAAGTCCGTCGATCGTGCACATCGTGCAGAGAAATTTCTCAAGTTGTGTCAT	300
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	301	CAACTCGGACTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGTGTGAC	360
Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140
Db	361	CGGCTCCACGACTATGACTGGACTTTTCGGCAGCGCTGGAAAGCTCGCAACCGGTGTG	420
Qy	141	ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle	160
Db	421	CGTGTGCGTACGGAAGCTGAATATGACGCTCGCAGCGATAGTGGCGCTGTGAAATC	480
Qy	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	481	ACTCAGGAAAGACGGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTCCAGTGTGAC	540
Qy	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200
Db	541	GTGAGTGAGAACATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGACTTCCAA	600
Qy	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220
Db	601	CTCGTATCTCTGGATTGCTGGCCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG	660
Qy	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal	240
Db	661	ATCGCGGTGTAAGCTCATCAAGTTCTCGTGGTATCTGACGAAATTTGCACTCGCGTT	720
Qy	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260
Db	721	GGTATGAGGTTGGTAGTGGCGAGGAAAGTGGTGCATGGCAACCGTGCACCCCTGG	780
Qy	261	GluIleAlaValAspAlaValGlyCysAspProGlnAlaLeuGluLeuTrpArgGluPhe	280
Db	781	GAATCCCTGTTGATGAGTGGCGGGATCCACACGGTTGGAACTGTGCGGAGAATTT	840
Qy	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	841	GAGTTTGGTTCGATGGGACGTCGGGCAATCGCGTGTTCGCTGGATTGCGTGGCCGAGCT	900
Qy	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal	320

Db	901	GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC	960
Qy	321	MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal	340
Db	961	ATGGTTCCGATCATTCGCGGCGCATCGTGATGATGATCGGACTTGTGGCGCTTACGTC	1020
Qy	341	PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis	360
Db	1021	TTGCGCGAGATCCTCGGACTCGTCCAAAGCTGGCGCGACTTGGGAAATCTTCGTGATCAC	1080
Qy	361	LeuHisTrpArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys	379
Db	1081	TTGATTTATCGATTGCCCGCAGCGATGTGCGGCGCCCGATAATATCGTTTCGAAG	1137
RESULT 2			
US-10-007-452-1			
; Sequence 1, Application US/10007452			
; Publication No. US20030093701A1			
; GENERAL INFORMATION:			
; APPLICANT: Tomb, Jean-Francois			
; APPLICANT: Bramucci, Michael G.			
; APPLICANT: Cheng, Qiong			
; APPLICANT: Kostichka, Kristy N.			
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors			
; FILE REFERENCE: C11709 US NA			
; CURRENT APPLICATION NUMBER: US/10/007,452			
; CURRENT FILING DATE: 2001-11-08			
; PRIOR APPLICATION NUMBER: 60/254,868			
; PRIOR FILING DATE: 2000-12-12			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 1			
; LENGTH: 1140			
; TYPE: DNA			
; ORGANISM: Rhodococcus AN12			
US-10-007-452-1			
Alignment Scores:			
Pred. No.:	8,33e-216	Length:	1140
Score:	2005.00	Matches:	379
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-10-007-527A-2 (1-379) x US-10-007-452-1 (1-1140)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	1	ATGACCGAGTAAAGTGTGAACACCTTTCGGCAAGACCGCGCTCCGCTCGTGTGCG	60
Qy	21	SerAspLysArgGlyLleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40
Db	61	TCCGATAGCCGCGCATCCGCGACGAACTGCGACCCAACTTCAACAAATCACCACGTCA	120
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	121	GAACATTTAACCGCTTGGCGCGCATTCGCGTGAACCGGTGTGACCATTTGTCAAC	180
Qy	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80
Db	181	GGTCCGAAAGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC	240
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100
Db	241	TGCTGTGCGGAAAGTCCGTCGATCGTGCACATCGTGCAGAGAAATTTCTCAAGTTGTCTCAT	300
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	301	CAACTCGGACTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGTGTGAC	360
Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140

Db 361 CGGCTCCAGACCTATGGACTTGGAGCTTCCGAGCCTGGAAAGCTGCGACCAACGGTCGT 420
 QY 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 Db 421 CGTTGGCGTACGGACGTGAATGTACGGCTGGACGGATACGTGCGCGTGTGAATC 480
 QY 161 ThrHisGlyLysAsnGlyTyrPheHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 Db 481 ACTCAGGAAAAAAGCGCTGGCAGCTCCAGTTTCACGGCTACTCATGTTTCAGTGGTAC 540
 QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
 Db 541 GTGAGTGAGAACATCTCCGAATCTTCTCGATGCGATGTTCGATCGGTGGACTTCCAA 600
 QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLysLeuAspValArgLys 220
 Db 601 CTGCTATCTCTGGATTGCTGGCCACACACGTAATTCGGGTGGTCTCGATGTACGAAAG 660
 QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 Db 661 ATCGCGGTGAAGCTCATCAAGTTCTCGCTGGTATCTGACGAAATTCGATCTGGCGTT 720
 QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
 Db 721 GGTATGGAGGTGGTAGTGGCGCAGGAAAAAGTGTGCGACATGGCAACCGTGCCACCTCG 780
 QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuLeuTyrArgGluPhe 280
 Db 781 GAATTCGCTGTTGATGCACTGGCGGGATCCACAGCGTTGGAACTGTGGCGAGAAATT 840
 QY 281 GluPheGlySerMetGlyArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
 Db 841 GAGTTTGGTTCGATGGAGCTGGGCAATCGCGTCCGTCGGTGGATGGTCCCGCAGCT 900
 QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 Db 901 GGTCTTGGCGGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAATCTTGGCGCTTACGTC 960
 QY 321 MetValAlaIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340
 Db 961 ATGGTTGCGATCATCTCCGCGCGATCGTGGATGATGATTCGGACTTGTGCGCTTACGTC 1020
 QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
 Db 1021 TTGGCGAGATCTCTGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC 1080
 QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
 Db 1081 TTGCATTATCGATTGCCGCGAGCGATGTCGGCGCCCGGATAATATATCGTTCCGAAG 1137

RESULT 3

US-10-415-562A-1

; Sequence 1, Application US/10415562A

; Publication No. US20040115661A1

; GENERAL INFORMATION:

; APPLICANT: E.I. du Pont De Nemours and Company

; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

; FILE REFERENCE: CL1709 US PCT

; CURRENT APPLICATION NUMBER: US/10/415.562A

; CURRENT FILING DATE: 2003-11-17

; PRIOR APPLICATION NUMBER: 60/254,868

; PRIOR FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1140

; TYPE: DNA

; ORGANISM: Rhodococcus AN12

US-10-415-562A-1

Alignment Scores:

Pred. No.: 8.33e-216 Length: 1140

Score: 2005.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-415-562A-1 (1-1140)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
 Db 1 ATGACCAGCGTAAAGTCTGACACACCTTCCGGCAAGACCGGCTCCCGCTCCGTGTCG 60
 QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
 Db 61 TCCGATAAGCGCGCATCCGGCACGAACCTCGCACCCAACTTCAACAAATCACCACGTCA 120
 QY 41 GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
 Db 121 GAAACATTTAAACCCCTGTGGCCGCGGATTTCTGGCGTGAACGCTGACCATGTGCAAC 180
 QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
 Db 181 GGTCCGAAAGGTTCTGGATTCCGAGGCCTTCGTTCTCGGAAAGGCGTGGATCTGCCCC 240
 QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
 Db 241 TGCTGTGGGAAAAAGTCCGTGCACATCGTCAGACGAAATTTCTCAAGTTGTGCTCAT 300
 QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 Db 301 CAACTCCGGACCTGGATCTGTTCGATGGTGCATGACCATCGGCATACAGCTGCTCAG 360
 QY 121 ArgLeuHisAspLeuTyrThrGlyLeuSerAlaAlaTyrLysAlaAlaThrAsnGlyArg 140
 Db 361 CGGCTCCAGACCTATGACTGGACTTTCGGCAGGCTGGAAAGCTGGCAACCGTGGT 420
 QY 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 Db 421 CGTTGGCGTACGGAACGTGAATGTACGGCTCGCAGCGGATACGTGCGCGCTGTGAAATC 480
 QY 161 ThrHisGlyLysAsnGlyTyrPheHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 Db 481 ACTCAGGAAAAAAGCGCTGGCAGCTCCACGTTTCACGCGCTACTCATGTTTCAGTGGTGC 540
 QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
 Db 541 GTGAGTGAGACATCTCGAATCTTCTCGAGATGGATGTTTCGATCGGTGGACTTCCAAA 600
 QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
 Db 601 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAAATTCGGGTGGTCTCGATGTACGAAAG 660
 QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 Db 661 ATCGCGCGTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTTGCACTTCGGGTT 720
 QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
 Db 721 GGTATGAGGTTGGTAGTGGCGACGGAAAAAGTGGTGCATGGAACCCGTGGACCCCTGG 780
 QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
 Db 781 GAAATCGCTGTGATGACGTGGCGGGATCCCAACAGCTTGGAACTGTGCGCGGAAATTT 840
 QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
 Db 841 GAGTTTGGTTCGATGGACGCTCGGCAATCCGCTGGTCCCGTGGATGGCTGCCCGAGCT 900
 QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluGluSerAlaProVal 320
 Db 901 GGTCTTGGGCGACAACTAACAGATGCTCAGATGTTTCAGATCGTGAGCAAGAAATCTGCCCGCGTC 960
 QY 321 MetValAlaIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340

Db	961	ATGTTTCGATCATTCGGCGCGATCGTGATGATGATTCGATTCGGCTTGTGGCCCTTACGTC	1020
Qy	341	PheGlyGluLeuLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis	360
Db	1021	TTCGGCGAGATCCCTCGGACTCGTCAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC	1080
Qy	361	LeuHisTyrArgLeuProAlaAlaAspValArgProProLeuLeuSerValArgLys	379
Db	1081	TTGCATTATCGATTGCCCGCAGCGATGTGCGGCGCCCGCAATATATCGGTTTCGCAAG	1137
RESULT 4			
US-10-007-527A-5/c			
; Sequence 5, Application US/10007527A			
; Publication No. US20030044807A1			
; GENERAL INFORMATION:			
; APPLICANT: Tomb, Jean-Francois			
; APPLICANT: Bramucci, Michael G.			
; APPLICANT: Cheng, Qiong			
; APPLICANT: Kostichka, Kristy N.			
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors			
; FILE REFERENCE: CL1709 US NA			
; CURRENT APPLICATION NUMBER: US/10/007,527A			
; CURRENT FILING DATE: 2001-12-05			
; PRIOR APPLICATION NUMBER: 60/254,868			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 5			
; LENGTH: 6334			
; TYPE: DNA			
; ORGANISM: Rhodococcus AN12			
US-10-007-527A-5			
Alignment Scores:			
Pred. No.:	7,46e-215	Length:	6334
Score:	2005.00	Matches:	379
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-10-007-527A-2 (1-379) x US-10-007-527A-5 (1-6334)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	3051	ATGACCAGCGTAACTGCTGAACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTCG	2992
Qy	21	SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer	40
Db	2991	TCCGATAAGCGCGCATCCGGCAGAACTCGCAGACCCCAAACTTCAACAAATCACACGTCA	2932
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	2931	GAAACATTAAACCTGTGGCGGCGGATTTCTGGCTGAACGGTGTGACCAATGTCTAAC	2872
Qy	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyThrIleCysPro	80
Db	2871	GGTCCGAAAGTTCTCGATTCCGAGGCGCTTCTCTCTCGGAAAGGGCTGGATCTGCCCC	2812
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValAlaHis	100
Db	2811	TGCTGTCCGGAAGTCCGGTGCACATCGTGACACGAAATTTCTCAAGTTGTGCTCAT	2752
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	2751	CAACTCCGGACTGATCTGTTCGATGTGTGACGATGACCATGCGCCATACAGCTGGTCAG	2692
Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140
Db	2691	CGGCTCCACGACCTATGACTGGACTTTCGGCAGCGTGGAAAGCTGCCACCAACGGTCTGT	2632
Qy	141	ArgTrpArgGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160

Db	2631	CGTTGGCGTACGGAACCGTGAATGTACGGCTCGACGATACGTGCGCGCTTGTGAATC	2572
Qy	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	2571	ACTCACGGAATAACCGCTGGCAGCTCCAGCTTCACGGCTACTCATGTTTCAGTGGTGAC	2512
Qy	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200
Db	2511	GTGAGTGAGAACATCTCTCGAATCTTCTCGATCGATGTTTCGATCGGTGGACTTCCAAA	2452
Qy	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220
Db	2451	CTCGTATCTCTGGGATTTCCTGCGCACACTAGCTAATTCGGGTGGTCTCGATGACGAAG	2392
Qy	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal	240
Db	2391	ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGTATCTGACGAAATTCATCTGCGCTT	2332
Qy	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260
Db	2331	GGTATGGAGGTTGGTAGTGGCGACGGAATAAGTGGTGCACATGGCAACCGTGCACCTGG	2272
Qy	261	GluIleAlaValAspAlaValGlyClyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280
Db	2271	GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAGCGTTTGAACACTGTGGCGAAATTT	2212
Qy	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	2211	GAGTTTGGTTCGATGGGACGTCGGCAATCGGCTGGTCCCCTGGATTCGCTGCCGAGCT	2152
Qy	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal	320
Db	2151	GGTCTTGGCGAGAACTAACAGATGCTCAGATCGTTGACAGAGAAATCTGCCCGGTC	2092
Qy	321	MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal	340
Db	2091	ATGGTTTCGATCATTCGGCGCGATCGTGGATGATGATTCGGACTTGTGGCCCTTACGTC	2032
Qy	341	PheGlyGluIleGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis	360
Db	2031	TTCCGGCGAGATCTCGGACTCGTCAGACTCGTCAGACTCGGCGACTTGGGAAATCTTCGTGATCAC	1972
Qy	361	LeuHisTyrArgLeuProAlaAlaAspValArgProProIleLeuSerValArgLys	379
Db	1971	TTGCAITATCGATTGCGCGCGATGTCGGCCCCCGATATATATATATATATATATATAT	1915
RESULT 5			
US-10-007-452-5/c			
; Sequence 5, Application US/10007452			
; Publication No. US20030093701A1			
; GENERAL INFORMATION:			
; APPLICANT: Tomb, Jean-Francois			
; APPLICANT: Bramucci, Michael G.			
; APPLICANT: Cheng, Qiong			
; APPLICANT: Kostichka, Kristy N.			
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors			
; FILE REFERENCE: CL1709 US NA			
; CURRENT APPLICATION NUMBER: US/10/007,452			
; CURRENT FILING DATE: 2001-11-08			
; PRIOR APPLICATION NUMBER: 60/254,868			
; PRIOR FILING DATE: 2000-12-12			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 5			
; LENGTH: 6334			
; TYPE: DNA			
; ORGANISM: Rhodococcus AN12			
US-10-007-452-5			
Alignment Scores:			
Pred. No.:	7,46e-215	Length:	6334
Score:	2005.00	Matches:	379
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-10-007-527A-2 (1-379) x US-10-007-527A-5 (1-6334)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	3051	ATGACCAGCGTAACTGCTGAACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTCG	2992
Qy	21	SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer	40
Db	2991	TCCGATAAGCGCGCATCCGGCAGAACTCGCAGACCCCAAACTTCAACAAATCACACGTCA	2932
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	2931	GAAACATTAAACCTGTGGCGGCGGATTTCTGGCTGAACGGTGTGACCAATGTCTAAC	2872
Qy	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyThrIleCysPro	80
Db	2871	GGTCCGAAAGTTCTCGATTCCGAGGCGCTTCTCTCTCGGAAAGGGCTGGATCTGCCCC	2812
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValAlaHis	100
Db	2811	TGCTGTCCGGAAGTCCGGTGCACATCGTGACACGAAATTTCTCAAGTTGTGCTCAT	2752
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	2751	CAACTCCGGACTGATCTGTTCGATGTGTGACGATGACCATGCGCCATACAGCTGGTCAG	2692
Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140
Db	2691	CGGCTCCACGACCTATGACTGGACTTTCGGCAGCGTGGAAAGCTGCCACCAACGGTCTGT	2632
Qy	141	ArgTrpArgGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160

Length: 6334
Matches: 379

7,46e-215
2005.00

Alignment Scores:
Pred. No.:
Score:

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-10-007-527A-2 (1-379) x US-10-007-452-5 (1-6334)			
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	3051	ATGACGAGCGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTG	2992
QY	21	SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer	40
Db	2991	TCCGATAAGCGCGCATCCGCGACGAACCTCGACACCAAACTTCAACAAATCACACGTC	2932
QY	41	GluThrPheAsnAlaCysGlyArgProLysGlyValAsnGlyValThrLeuValAsn	60
Db	2931	GAACAATTTAACCCCTGTGCGCGCGGATTTCTGGCGTGAACGCGTGCACATTGTCAAC	2872
QY	61	GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpLysCysPro	80
Db	2871	GGTCCGAAGGTTCTCGATTCCGAGGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC	2812
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValAlaHis	100
Db	2811	TGCTGTGCGGAAAGTGTGCGATGCGATCGTCAGACGAAATTTCTCAAGTTGTCTCAT	2752
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	2751	CAACTCGGACGACTGATCTGTTGCGATGGTGAACGATGACCATGCGCATACAGTGTCTAG	2692
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140
Db	2691	CGGCTCCACACCATATGACTGGACTTTCGCGACGCTTGGAAAGCTGCGACCAACGTCGT	2632
QY	141	ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLeu	160
Db	2631	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGAGGATACGTGCGCGCTGTGAAATC	2572
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	2571	ACTCAGGAATAAAGCGCTGGCACGTCACGTTCCAGCGCTACTCATGTTGAGTGTGAC	2512
QY	181	ValSerGluAsnLeuLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200
Db	2511	GTGAGTGAGAACATCTCTGATCTCTCGATGCGATGTCGATCGGTGGACTTCCAAA	2452
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLysLeuAspValArgLys	220
Db	2451	CTCGTATCTCTGGGATTTGTCGCGCACTACGTAATTCGGGTGCTCTCGATGTACGAAG	2392
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal	240
Db	2391	ATCGCGCGTGAGCTGATCAAGTTCTCGCTCGATCTGACGAAATTTGATCTCGCGTT	2332
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260
Db	2331	GGTATGAGGTTGTTAGTGGCGACGGAAGTGTGACATGCAACCGTGCACCTCGG	2272
QY	261	GluIleAlaValAspAlaValGlyLysProGlnAlaLeuGluLeuTrpArgGluPhe	280
Db	2271	GAATTCGCTCTGATGAGTGGCGGGATCCACAGCGTTGGAACCTGGCGGAGATTT	2212
QY	281	GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	2211	GAGTTTGTTGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTGCGTGCACGCT	2152
QY	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluLeuSerAlaProVal	320
Db	2151	GGTCTTGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC	2092
QY	321	MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal	340
Db			

RESULT 6

US-10-415-562A-5/c
 ; Sequence 5, Application US/10415562A
 ; Publication No. US20040115661A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont De Nemours and Company
 ; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
 ; FILE REFERENCE: CL1709 US PCT
 ; CURRENT APPLICATION NUMBER: US/10/415,562A
 ; PRIOR FILING DATE: 2003-11-17
 ; PRIOR APPLICATION NUMBER: 60/254,868
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 5
 ; LENGTH: 6334
 ; TYPE: DNA
 ; ORGANISM: Rhodococcus AN12
 US-10-415-562A-5

Alignment Scores:

Pred. No.: 7,46e-215 Length: 6334
 Score: 2005.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-415-562A-5 (1-6334)

QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	3051	ATGACGAGCGTAAGTGTGAACACCTTTCGCGCAAGACCGGCTCCCGTCTCGTGTG	2992
QY	21	SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnThrThrSer	40
Db	2991	TCCGATAAGCGCGCATCCGCGACGACGACCCAACTTCAACAAATCACACGTC	2932
QY	41	GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn	60
Db	2931	GAACAATTTAACGCTGTGCGCGCGGATTTCTGGCGTGAACGCTGTGACCATGTCAAC	2872
QY	61	GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpLysCysPro	80
Db	2871	GGTCCGAAGGTTCTGATTCGGAGGCTTGTTCCTCGGAAAGGCTGGATCTGCCCC	2812
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValAlaHis	100
Db	2811	TGCTGTGCGGAAAGTGTGCTGACATCTGTCGACGAAATTTCTCAAGTTGTTCAT	2752
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	2751	CAACTCGGACGACTGATCTGTTGCGATGGTGAACGATGACCATGCGCCATACAGTGTGTCAG	2692
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140
Db	2691	CGGCTCCACACCATATGACTGGACTTTCGCGACGCTTGGAAAGCTGCGACCAACGTCGT	2632
QY	141	ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLeu	160
Db	2631	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGAGGATACGTGCGCGCTGTGAAATC	2572
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisAlaLeuLeuMetPheSerGlyAsp	180

Db 2571 ACTCAGGAAAAACCGGTGCGCTCCACGTTCCACGCGTACTCATGTTTCAGTGGTGAC 2512
Qy 181 ValSerGluAenIleLeuGluSerPheSerAspAlaMetPheAspArgTrrThrSerlys 200
Db 2511 GTGAGTGAGAACATCCTCGAATCCTTCTCGGATCGGATGTTTCGATCGGTGACITCCAA 2452
Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArglys 220
Db 2451 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAAATCGGTGGTCTCGATGTACGAAG 2392
Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaIleAlaIleAlaIleAlaSerGlyVal 240
Db 2391 ATCGCGGTGAGAGTGTGATCAAGATTTCTCGCTCGTATCTGACGAAATTTGCAATCTGGCGTT 2332
Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 2331 GGTATGGAGGTTGGTAGTGGCGACGGAAGAGTGGTCCACATGGCAAACCGTGACCCCTGG 2272
Qy 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 2271 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAAAGCGTTTGGAACTGTGGCGAGAATTT 2212
Qy 281 GluPheGlySerMetGlyArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 300
Db 2211 GAGTTTGGTTCGATGGGACGTCGGCAATCGCGTGTCCCGTGGATTTGGTCCCGGAGCT 2152
Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 2151 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGACAGGAAGATCTGCCCGGTC 2092
Qy 321 MetValAlaIleIleProAlaArgSerTrpMetIleArgThrCysAlaProTrpVal 340
Db 2091 ATGGTTGGCATCATTCGGCGCGATCGTGGATGATGATTCGGACTTTGTGGCGCTTACGTC 2032
Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaIleAlaIleAlaIleAlaIleAlaIle 360
Db 2031 TTCCGCGAGATCCTCGACTCGTCGAACTGGCGGACCTGGGAAATCTTCTGTGATCAC 1972
Qy 361 LeuHisTrpArgLeuProAlaAspValArgProProIleSerValArglys 379
Db 1971 TTGCATATCGATTGCGCGAGCGGATGTGGGCCCCCGGATATATATCGGTTCGCAAG 1915

RESULT 7
US-10-007-527A-7/c
; Sequence 7, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-527A-7

Alignment Scores:
Pred. No.: 1.28e-214 Length: 9652
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-007-527A-7 (1-9652)
Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
Db 5052 ATGACCAGCGTAAGTGTGAACACCTTTCCGGCAAGACCCGGCTCTCGCTCTGCTCG 4993
Qy 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
Db 4992 TCCGATTAAGCCCGCATCCGCGACGAACTGCGACCCCAAACTTCAACAAATCACCACGTCA 4933
Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 4932 GAAACATTTAAACGCTGTGGCGCGCGGATTTCTGGCGTGAACGGTGTACCATTTGTCAAC 4873
Qy 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db 4872 GGTCCGAAAGGTTCTGGATTCGGAGGCCCTTCTCTCCGCGGAAAGGGCTGGATCTGCCCC 4813
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 4812 TGCTGTGCGGAAAGTCTGGTGCACATCTGCAGACGAAATTTCTCAAGTTTGTGCTCAT 4753
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 4752 CAACTCCGAGCTGGATCTGTTGCGATGGTGCAGATGACCATCGCCCATACAGCTGGTCAG 4693
Qy 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 4692 CGGCTCCACGACTATGAGCTTTCGGCAGCTCGGAAAGCTTGGAAAGCTTGTGAAATC 4633
Qy 141 ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle 160
Db 4632 CGTTGGCTGACGGAACGTGAATGTACGCTGCGAGGATACGTGGCGCTGTGAAATC 4573
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 4572 ACTCAGGAAAAACCGCTGGCAGCTCCACGTTACGCGCTACTCATGTTTCAGTGGTGAC 4513
Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 4512 GTGAGTGAACCAATCTCGAATCCTTCTCGATGCGATGTTTCGATCGGTGGACTTCCAA 4453
Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 4452 CTCGTATCTCTGGATTTGCTGGCCACTACGTAAATTCGGTGGTCTCGATGTACGAAAG 4393
Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 240
Db 4392 ATCGCGGTGAAGCTGATCAAGTTCCTCGCTGGCTATCTGACGAAATTTGATCTGGCGTT 4333
Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 4332 GGTATGGAGGTTGGTAGTGGCGGACGGAAGGTTGTCAGATCGCAACCGTGGACCTGG 4273
Qy 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 4272 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAAAGCTTGGAACTGTGGCGAGAATTT 4213
Qy 281 GluPheGlySerMetGlyArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 300
Db 4212 GAGTTTGGTTCGATGGGACGTCGGCAATCGGTGGTTCGGTGGATTTGGCTGGCCGAGCT 4153
Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 4152 GGTCTTGGGCGAAGCAATCAAGATGCTCAGATCGTTGAGCAGGAAGATCTGCCCGGTC 4093
Qy 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
Db 4092 ATGGTTGGATCATTTCCGGCGCGATCGTGGATGATGATTCGCACTTGTGGCGCTTACGTC 4033
Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaIleAlaIleAlaIleAlaIleAlaIle 360

Db 4032 TTCGGCGAGATCTCGGACTCGTCGAAGTCGGCGACTTGGGAAAATCTTCGTGATCAC 3973
QY 361 LeuHistiYrArgLeuProAlaAlaAspValArgProProleuSerValArgLys 379
|||||
Db 3972 TTGCATTATCGATTGCCCGAGCGATGTCGGCCCGGATATATATCGGTTCGCAAG 3916
RESULT 8
US-10-007-452-7/c
; Sequence 7, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; FILE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-452-7
Alignment Scores:
Pred. No.: 1,28e-214 Length: 9652
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-007-527A-2 (1-379) x US-10-007-452-7 (1-9652)
QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
|||||
Db 5052 ATGACCGGTAAGTGCCTGGAACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTG 4993
QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGluThrThrSer 40
|||||
Db 4992 TCCGATAAGCGGCATCGGCACGAATCGGACCACTTCCGCAAACTTCAACAATCACCACGTCA 4933
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
|||||
Db 4932 GAAACATTTAAGCGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCAC 4873
QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80
|||||
Db 4872 GGTCCGAAGGTTCTGGAATCGGAGGCTTCTGTTCTCCGGAAGGGCTGGATCTGCCCC 4813
QY 81 CysCysAlaGlyLysValGlyValAlaHisArgAlaAspGluLeuSerGlnValValAlaHis 100
|||||
Db 4812 TGCTGTGGGAAAGTCGGTGCACATCGTCGACAGCAAAATTTCTCAAGTTGTGTGCTCAT 4753
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
|||||
Db 4752 CAATCGGGACTGGATCTGTTCGATGGTGAAGATGACCATGCGCCATACAGCTGGTGCAG 4693
QY 121 ArgLeuHisAspLeuTyrThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
|||||
Db 4692 CGGCTCCACACCTATGACTGGACTTTCGCGAGCTTTCGCGAGCTGGAAAGCTGGCAACCGTGT 4633
QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
|||||
Db 4632 CGTTGGCGTACGNAAGTGAATGATGCTGCGGACGAGTACGTGCGGCTGTGTGAATC 4573
QY 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisValHisValHisValHisValHis 180
|||||

Db 4572 ACTCAGGAAAAAAGCGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTTCAGTGTGAC 4513
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
|||||
Db 4512 GTGAGTGAGAACATCTCGAATCTTCTCGATGCGATGTTTCGATCGTGGACTTCCAAA 4453
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
|||||
Db 4452 CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAAATTCGGGTGCTCTCGATGTACGAAG 4393
QY 221 IleGlyGlyGluAlaAspGlnValIleAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
|||||
Db 4392 ATCGCGGTGAAGCTGATCAAGTTCTCGCTGCGATCTCACGAAAAATTCGATCTGGCGTT 4333
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
|||||
Db 4332 GGTATGAGGTTGTTAGTGGCGACGGAAGTGTGCGATGCGCAACCGTGACCTGG 4273
QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
|||||
Db 4272 GAAATCGCTTTGATGACGTGGCGGATCCACAAGCTTGGAACTGTGGCGAGATTT 4213
QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
|||||
Db 4212 GAGTTTGGTTTCATGGGACGTGGGCAATCGCGTGTGCGATTTGCGTCCGAGCT 4153
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
|||||
Db 4152 GGTCTTGGCGAGAACTAACAGATCTCAGATCTGTGAGCAGAAAGAAATCTGCCCGGTC 4093
QY 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
|||||
Db 4092 ATGTTTGGATCATTCGCGCGCATCGTGGATGATGATTCGGACTTGTGGCTTACGTC 4033
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
|||||
Db 4032 TTCGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC 3973
QY 361 LeuHistiYrArgLeuProAlaAlaAspValArgProProleuSerValArgLys 379
|||||
Db 3972 TTGCATTATCGATTGCCCGAGCGATGTCGGCCCGGATAATATATCGTTTCGCAAG 3916

RESULT 9

US-10-415-562A-7/c
; Sequence 7, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-415-562A-7

Alignment Scores:

Pred. No.: 1,28e-214 Length: 9652
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-415-562A-7 (1-9652)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

```
Db 5052 ATGACGAGCGTAAAGTGTGAACACCTTTCCGGCAAGACCGCGCTCCGCTCGTGTGC 4993
Qy 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
Db 4992 TCCGATTAAGCGCGGCAATCCGACGAACTGCGACCAAACTTCAACAAATCACACGTC 4933
Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 4932 GAAACATTTAAACGCTGTGGCGCGCGGATTTCTGGCGTGAACGGGTGACCATTTGTCAAC 4873
Qy 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db 4872 GGTCCGAAGGTTCTGGATTCTGGAGGCTTCTGCTCCGGAAGAGGCTGGATCTGCCCC 4813
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 4812 TGTGTGGCGGAAAGTCTGGTGCACATCTGTCAGACGAAATTTCTCAAGTTGTGCTCAT 4753
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 4752 CAACCTCGGAGCTGGATCTGTGGATGTGTGACGATGACCATCGCCATACACGCTGGTCAG 4693
Qy 121 ArgIleuHisAspLeuThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 4692 CGGCTCCACGACCTATGGACTTTCGGAGCTTTCGGCAGCTGGAAAGCTCGGACCAACGGTCTGT 4633
Qy 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 4632 CTTTGGCGTACGGAACGTGAATGTACGGCTCGGACGGATACGTGGCGCTGTGAAATC 4573
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 4572 ACTCACGGAAAAACCGCTGGACGCTCCACGTTCCAGCGCTACTCATGTTCAGTGTGTGAC 4513
Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 4512 GTGAGTGAAGAAATCTCGAATCTTCTCGATTCGATGTTTGAATCGGTGGATCTCCAAA 4453
Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyCysValArgLys 220
Db 4452 CTCGTATCTCTGGATTTGCTGGCCACTACGTAATTCGGTGTCTCGATGTACGAAAG 4393
Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrIleuThrLysIleAlaSerGlyVal 240
Db 4392 ATCGCGGCTGAAGCTCATCAAGTCTCTCGTGGTATCTGACGAAATTTGATCTGGCGTT 4333
Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 4332 GGTATGGAGTTGTTAGTGGCGACGGAAGAAAGTGTGACATGCGCAACCGTGCACCTGG 4273
Qy 261 GluIleAlaValAspAlaValGlyArgProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 4272 GAAATCGCTGTTGATGAGTGGCGGGATCCAAAGCGTTGGAACCTGTGGCAGAAATTT 4213
Qy 281 GluPheGlySerMetGlyArgAlaAlaIleAlaTrpSerArgGlyIleuArgAlaProAla 300
Db 4212 GAGTTTGGTTTCGATGGGACGTCGGCAATCGGTGGTCCGTGGATGGATGGTGGCGGAGCT 4153
Qy 301 GlyLeuGlyValGluLeuThrAspAlaGlnIleValGluGlnGluLeuSerAlaProVal 320
Db 4152 GGTCTTGGGCGAAGATPACAGATGCTCAGATGTTGAGCAGAGAGAAATCTGCCCGGTC 4093
Qy 321 MetValAlaIleIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db 4092 ATGGTTGCGATCATTTCCGGCGCGATCTGTGATGATGATTCGAACTTGTGGCGCTTACGTC 4033
Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 4032 TTCGGCGAGATCTCTCGACTCTGTCAAGCTGTGGCGGACTTGGGAAATCTTCGGATCAC 3973
Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProIleIleSerValArgLys 379
```

```
Db 3972 TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCCCCCGGATAATATCGTTGCGCAAG 3916
RESULT 10
US-10-007-527A-6/c
; Sequence 6, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
; US-10-007-527A-6
Alignment Scores:
Pred. No.: 1,55e-214 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1.4 Gaps: 0
US-10-007-527A-2 (1-379) x US-10-007-527A-6 (1-11241)
Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
Db 6641 ATGACGAGCGTAAAGTGTGAACACCTTTCCGGCAAGACCGCGCTCCGCTCGTGTGC 5582
Qy 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
Db 6581 TCCGATTAAGCGCGGCAATCCGACGAACTTCAACAAATCACACGTC 6522
Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 6521 GAAACATTTAAACGCTGTGGCGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 6462
Qy 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db 6461 GGTCCGAAAGGTTCTGGATTCTGGAGGCTTCTGCTTCCGGAAGGCGTGGATCTGCCCC 6402
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 6401 TGTGTGGCGGAAAGTCTGGTGCACATCTGTCAGACGAAATTTCTCAAGTTGTGCTCAT 6342
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 6341 CAACCTCGGAGCTGGATCTGTGGATGTGTGACGATGACCATGCCATACACGCTGGTCAG 6282
Qy 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 6281 CGGCTCCACGACCTATGGACTTTCGGAGCTTTCGGAGCTTGGAAAGCTGGCAACCGTCTGT 6222
Qy 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 6221 CGTTGGCGTACGGAACCTGGAATGTACGGTGCACGATACGTCGCGCTGTGTAATC 6162
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 6161 ACTCACGGAAAAACCGCTGGCAGCTCCAGCTTCCACGCGCTACTCATGTTCAGTGTGAC 6102
Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
```

Db 6101 GTGAGTGGAGAACATCTCTCGAATCTTCTCGGATGCGATGTTTCGATCGGTGGACTTCCAAA 6042
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 6041 CTGCTATCTCTGGATTGTCGCGCACTACGTAAATTCGGGTGTTCTCGATGTACGAAAG 5982
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 5981 ATCGCGGTGAAGCTGATCAAGTCTCTCGTTCGTAATTCGACGAAATTCGATCTGGGTT 5922
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 5921 GGTATGGAGGTGGTAGTCGACGCGAAAGTGGTCGACATGCCAACCGTGCACCTGG 5862
QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 5861 GAAATCGCTGTGATCGATGGCGGGGATCCAAAGCGTTGGAATCTGGCGGAGAATTT 5802
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 5801 GAGTTTGGTTCGATGGAGCTCGGGCAATCGGTGGTCCGTGGATGGTCCCGGAGCT 5742
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 5741 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGACGAGCAAGAAATCTGCCCGGTC 5682
QY 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db 5681 ATGGTTGCGATCATTCGCGCGGATCGTGATGATGATTCGGACTTGTGCGCTTACGTC 5622
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 5621 TTGCGCGAGATCCTCGGACTCGTGAAGCTGCGGAGCTTGGGAAATCTTCGTGATCAC 5562
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleSerValArgLys 379
Db 5561 TTGCATTATCGATTGCCGCGAGGATGTGCGGCCCGCGATAATATCGTTTCGCAAG 5505

RESULT 11

US-10-007-452-6/c
; Sequence 6, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CLI709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-452-6

Alignment Scores:

Pred. No.: 1,55e-214 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-007-452-6 (1-11241)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
|||||

Db 6641 ATGACCAGCGTAAGTGTCTGAACACCTTTCGCGCAAGACCGGCTCTCCGTCCTCTGTGTCG 6582
QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
Db 6581 TCCGATAAGCGCGCATCCGGCACGAACCTCGACCCCAAACTTCAACAAATTCACACGTCAC 6522
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 6521 GAAACATTTAAACCCCTGTGCGCGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 6462
QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db 6461 GGTCCGAAAGTTCCTCGATTGCGAGGCTTCTGCTCGGAAAGGCGTGGATCTGCCCC 6402
QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 6401 TGCTGTGCGGAAAAGTGGTGCACATCTGTCAGACGAAATTTCTCAAGTTGTGTCTCAT 6342
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 6341 CAACTCGGAGCTGGATCTCTGTGCGATGGTGACGATGACCATGCGCCATACAGCTGGTCAG 6282
QY 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 6281 CGGCTCCAGACCTATGGACTGGACTTTCGCGAGCTCGAAAGCTCGGACCAACCGTCTGT 6222
QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 6221 CGTGGCGTACGGAACGTGAAATGTACGGCTCGACGCGATACGTGGCGCTGTGTAATTC 6162
QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 6161 ACTCAAGAAAGAAACCGGCTGGCACGTTCCACGTTCCACGCGTACTCATGTTCAAGTGGTAC 6102
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 6101 GTGAGTGAGAACATCCTCGAATCTTCTCGGATGCGATGTTTCGATCGGTGGACTTCCAAA 6042
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 6041 CTCGTATCTCTGGGATTTGCTCGGCCACTAGTAATTCGGTGGTGTCTCGATGTACGAAAG 5982
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 5981 ATCGCGGTGAAGCTGATCAAGTTCCTCGTGGTATCTGACGAAATTTGCAATCTCGGTT 5922
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 5921 GGTATGGAGGTGGTAGTGGCGACGCGAAAGTGGTCGACATGGCAACCGTGCACCTGG 5862
QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 5861 GAAATCGCTGTGATGCGAGTGGCGGGGATCCACAGCTTGGAACTGTGCGGAGAATTT 5802
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 5801 GAGTTTGGTTCGATGGAGCTCGGCAATCGGTGGTCCGTTGGATTTGCGTCCCGAGCT 5742
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 5741 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGACGAGCAAGAAATCTTCGTGATCAC 5682
QY 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db 5681 ATGGTTGCGATCATTCGCGCGGATCGTGATGATGATTCGGACTTGTGCGCTTACGTC 5622
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 5621 TTGCGCGAGATCCTCGGACTCGTGAAGCTGCGGAGCTTGGGAAATCTTCGTGATCAC 5562
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleSerValArgLys 379
Db 5561 TTGCATTATCGATTGCCGCGAGGATGTGCGGCCCGCGATAATATCGTTTCGCAAG 5505


```
Db 1485 GCGCGCGGAACGTCGTGGTCTCTG-----GGCTGATGCGG 1450
Qy 73 CysGlyLysGlyTyrPheCysProCysCysAlaGlyValGlyAlaHisArgAlaAsp 92
Db 1449 TGGCGCGGATCTGGCTCTCCCGGCTGCGCGCGCCACGATCCGGCACAAGCGCGCGAG 1390
Qy 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAla---MetValThr 111
Db 1389 GAGATCACCGCCCGCGTGGTGGATGATCAAGCGCGGGGGGACCGCTACCTGGTCACC 1330
Qy 112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrPheGlyLeuSerAla 131
Db 1329 TTCACGGCGCGCATGGGACACAGCGACCGCTCGCGGACCTCATGGACGCCCTCCAGGGC 1270
Qy 132 AlaTrpLysAlaAlaThrAsnGlyArg-----141
Db 1269 ACCCGAAGACGCGGACAGCCCCCGCGCGCGCTTACGACGACGTGATCAAGGGC 1210
Qy 142 -----TrpArgThrGlu-----145
Db 1209 GGCACGTGGCGCGGACCGCGCGCCCAAGACCGGGCCCGCGCGCGCGGCGATC 1150
Qy 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLys---164
Db 1149 CGAGACCGGATCGGGTACGTCGGCATGTCGCGCGCGCCAGGTCACCGTGGCGCAGATC 1090
Qy 165 AsnGlyTyrHisValHisAlaLeuLeuMetPheSerGlyAspVal-----181
Db 1089 AACGGCTGGCACCGCGACATCACCGATCGTCTGGTGGCGCGCGCGGACCGAGGGGAG 1030
Qy 182 -----SerGluAsnIleLeuGluSerPhe-----141
Db 1029 CGGTCCCGAAGACGATGTCGCGCACCTTCGAGCCGCGCGCGCGCGCGGATG 970
Qy 198 ThrSerLysLeuValSerLeuGlyPheAlaProLeuArg-----211
Db 969 CAGGGGCACTGGCGGTCGCGT-----TGGACCGCGCGCGCGCGCGCGCGCTTC 913
Qy 212 -----AsnSerGlyLeuAspValArgLysIleGlyGluAlaAsp---Gln 227
Db 912 ACGCCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 853
Qy 228 ValLeuAlaAlaTyrLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCGGAGTACATCGCCAAAGCCAGACCGCGCGCGCGCGCGCGCGCGCG 793
Qy 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle-----262
Db 792 CGCGCGGACCTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 733
Qy 263 AlaValAspAlaValGlyGly-----AspProGlnAla-----273
Db 732 ATCGGGGACCTGACCGCGCGCATGACCGAGACGACCGCGCGCGCGCGCGCGCG 673
Qy 274 -----LeuGluLeuTrpArgGluPheGlySerMetGlyArgArgAlaIleAla 291
Db 672 TGGAACTCTCGCGTGGCAGAGTACGAGCGGGCAACCGCGGGGACCGCGCGCG 613
Qy 292 TrpSerArgGlyLeuArgAlaAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCGCGCTACCTGCGCGAGATGCTCGGGCTCGACGCGCGCGCGCGCGCG 559
RESULT 14
US-09-754-112A-1/c
; Sequence 1, Application US/09754112A
; Publication No. US20020090687A1
; GENERAL INFORMATION:
; APPLICANT: TRIPATHI, Gyanendra, et al.
; TITLE OF INVENTION: Process for the Production of Polyhydroxyoctanoate by Streptomyces
; TITLE OF INVENTION: lividans
; FILE REFERENCE: 041144.006
; CURRENT APPLICATION NUMBER: US/09/754,112A
; CURRENT FILING DATE: 2001-01-05
```

```
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 12738
; TYPE: DNA
; ORGANISM: Streptomyces lividans
US-09-754-112A-1
Alignment Scores:
Pred. No.: 1,15e-26 Length: 12738
Score: 337.00 Matches: 107
Percent Similarity: 42.46% Conservative: 45
Best Local Similarity: 29.89% Mismatches: 126
Query Match: 16.81% Indels: 80
DB: Gaps: 16
US-10-007-527A-2 (1-379) x US-09-754-112A-1 (1-12738)
Qy 23 LysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGluThr 42
Db 1605 CGGAAGCGCGCGCGGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1546
Qy 43 PheAsnAlaCysGlyArg-----ProIleSerGlyVal-----53
Db 1545 TCACAGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1486
Qy 54 ---AsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyLeuArgSer 72
Db 1485 GCGCGCGGAACGTCGTGGTCTG-----GGCTGATGCGG 1450
Qy 73 CysGlyLysGlyTyrPheCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1449 TGGCGCGGATCTGGCTCTCCCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGAG 1390
Qy 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAla---MetValThr 111
Db 1389 GAGATCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1330
Qy 112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAla 131
Db 1329 TTCACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1270
Qy 132 AlaTrpLysAlaAlaThrAsnGlyArg-----141
Db 1269 ACCCGAAGACGCGCGACAGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1210
Qy 142 -----TrpArgThrGlu-----145
Db 1209 GGCACGTGGCGCGGACCGCGCGCCCAAGACCGGGCCCGCGCGCGCGCGCGATC 1150
Qy 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLys---164
Db 1149 CGAGACCGGATCGGGTACGTCGGCATGTCGCGCGCGCCAGGTCACCGTGGCGCAGATC 1090
Qy 165 AsnGlyTyrHisValHisAlaLeuLeuMetPheSerGlyAspVal-----181
Db 1089 AACGGCTGGCACCGCGACATCACCGATCGTCTGGTGGCGCGCGCGGACCGAGGGGAG 1030
Qy 182 -----SerGluAsnIleLeuGluSerPhe-----SerAspAlaMetPheAspArgTrp 197
Db 1029 CGGTCCCGAAGACGATGTCGCGCACCTTCGAGCCGCGCGCGCGCGCGGATG 970
Qy 198 ThrSerLysLeuValSerLeuGlyPheAlaProLeuArg-----211
Db 969 CAGGGGCACTGGCGGTCGCGT-----TGGACCGCGCGCGCGCGCGCGCGCTTC 913
Qy 212 -----AsnSerGlyLeuAspValArgLysIleGlyGluAlaAsp---Gln 227
Db 912 ACGCCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 853
Qy 228 ValLeuAlaAlaTyrLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCGGAGTACATCGCCAAAGCCAGACCGCGCGCGCGCGCGCGCGCGCG 793
```

QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle----- 262
Db 792 CGCGCCGACCTCAAGACGGCGGAGACGTCGCCCGGTTTCGAACCTCCTCGGACGG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla----- 273
Db 732 ATCGGGGACCTGACCGCGGATGACCGAGACGACGCGCGGGTCTCGCTGGAG 673
QY 274 -----LeuGluLeuTrpArgGluPheGlySerMetGlyArgAlaIleAla 291
Db 672 TGGAACTCTCGCGTGCACAGATACGAGCGGGCAACCGCGGAGCGCGGCCATCGAA 613
QY 292 TrpSerArgGlyLeuArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCCGCTACCTGCGGCAGATGCTCGGGCTCGACGCGGCGGACACCGAGGCC 559

RESULT 15

US-09-835-381-5
; Sequence 5, Application US/09835381
; Patent No. US20020045223A1
; GENERAL INFORMATION:
; APPLICANT: SUGA, Mikiko
; APPLICANT: ASAKURA, Yoko
; APPLICANT: MORI, Yukio
; APPLICANT: ITO, Hisao
; APPLICANT: KURAHASHI, Osamu
; TITLE OF INVENTION: ARGinine REPRESSOR DEFICIENT STRAIN OF CORYNEFORM BACTERIUM AND METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 206018US0
; CURRENT APPLICATION NUMBER: US/09/835,381
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: JP2000-129167
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 4447
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1318)..(2598)
; OTHER INFORMATION:
US-09-835-381-5

Alignment Scores:
Pred. No.: 5, 19e-26 Length: 4447
Score: 326.00 Matches: 101
Percent Similarity: 47.75% Conservative: 58
Best Local Similarity: 30.33% Mismatches: 133
Query Match: 16.26% Indels: 42
DB: 9 Caps: 14

US-10-007-527A-2 (1-379) x US-09-835-381-5 (1-4447)

QY 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
Db 1396 ATGTACAAGATCAACCAATAGTAAGGCTCTGGCGGGGTGCCATAGGTG-GCGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyLeuAr 71
Db 1455 AGCTGTGGGTGTCTCTGGTCTGCTACGGT-----GCTTCCGATTTGAGGGTCTGCA 1508
QY 71 gSerCysGlyLysGlyTrpIleCysProCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AAACCTCACTCTCGCTGGGGGTCACTCTGGTGAATTCGAAGTCAATGGCGGAACGCCG 1568
QY 91 AspGluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATTGAGCTGGCTATTCTACTAAGAATCACTTGGCGCGGGTGGCGCGTCTCAGTATT 1628
QY 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130

Db 1629 TGTGGGCACTGTTTCGACACCAACCGCTCACAGTTCATTTGCGCAGGTTGAAGCGGTATTAA 1688
QY 130 rAlaAlaTrpIlySalAlaAlaThrAsnGlyArgTrpArgThrGluArgGluMetTrpGl 150
Db 1689 GACTGCGTACTCTTCGATGTGTGAACCATCTCAGTGAAGAAAGAACTGCACCGTACGG 1748
QY 150 yCysAspGlyTrpValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa 169
Db 1749 GGTGGGACACACCTATAGTACATAGGTACACAGACTCTTTGGCGCAACGGTTGGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 1809 GCACCGCAACATGCTGTGTTCTTGGATCGTCCACTGTCTCAGCATGAACCTCAAGGCCGTT 1868
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCCATGTTTCCCGCTGTCGCTGGTGTGTTAAGCCGGTATGGACGCCGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGlyGluAlaAspGl 227
Db 1929 ACTGGTGACGACCGGGTCAAACTTGATCAGGTGTCTACCTGGGGTGGAGACGCTGCCGA 1988
QY 227 nValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGGCTC 2036
QY 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
Db 2037 CGCTACTAAACCCGCGCTCTAAGGGGTCTGACACGCGGTTTCAGATG---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
Db 2094 GGCCGATCAAGCGGCGCGGAGGATATGGACGCTGTTTGGTGGCTCGGTCGGCTGCA 2153
QY 279 uPheGluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
Db 2154 GTATGAGGTGGTCTTCTAAAAACCTCGTTCG---TCCTGGTCACTGGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu--- 315
Db 2205 GCGTGTCTTGGGCACTTGATTACATAGACCTGATGACGCTGTAATATGAGAGAACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGGTCTGGAAGCACCGGTCGAATCAACCCGCGTGTGCTGTGC 2324
QY 323 aIleIleProAlaArgSerTrpMetMetIleArgThr 335
Db 2325 TTTGGTGAAGCCGCGATGATTGGAAACTGATTCACTCT 2361

Search completed: October 23, 2004, 08:18:11
Job time : 659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 03:42:56 ; Search time 24 Seconds
(without alignments)
1519.423 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGKDRPPVLVS.....HLHYRLPAADVRPPIISVRK 379
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: Piri: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321.5	16.0	256	2	S34850
2	316	15.8	456	2	A31844
3	230.5	11.5	437	2	S04020
4	169	8.4	240	2	I40335
5	127	6.3	314	2	B35390
6	123.5	6.2	315	2	S25721
7	122.5	6.1	315	2	T44800
8	118.5	5.9	315	2	JN0856
9	115	5.7	336	2	A40651
10	114	5.7	314	2	T46831
11	111.5	5.6	333	2	A49783
12	105	5.2	381	2	A83561
13	104.5	5.2	340	2	I40561
14	104	5.2	420	2	AG2059
15	102.5	5.1	420	2	AC3284
16	101.5	5.1	339	2	S01038
17	100.5	5.0	1443	2	G75393
18	99.5	5.0	614	2	D87410
19	98.5	4.9	532	2	H83493
20	98.5	4.9	3535	2	B83641
21	98	4.9	442	2	D87594
22	95.5	4.8	606	2	D83484
23	94.5	4.7	1198	2	T28678
24	94.5	4.7	1763	2	T17465
25	94	4.7	317	2	A43621
26	94	4.7	382	2	AE2861
27	94	4.7	382	2	D97638
28	94	4.7	544	1	BVYCGL
29	94	4.7	1245	2	H87340

30 93 4.6 317 2 S25294
31 93 4.6 748 2 T37097
32 93 4.6 736 2 S57844
33 92.5 4.6 354 2 B75355
34 92.5 4.6 600 2 E72027
35 92.5 4.6 600 2 A85597
36 92.5 4.6 1026 2 G87346
37 92 4.6 799 2 T48889
38 91.5 4.6 302 2 H96792
39 91.5 4.6 332 2 S51694
40 91.5 4.6 350 2 B82281
41 91 4.5 296 1 OPNE7
42 91 4.5 356 2 B83200
43 90.5 4.5 320 2 D96750
44 90.5 4.5 2399 2 AH3009
45 90.5 4.5 2399 2 F98274

ALIGNMENTS

RESULT 1

S34850
hypothetical protein - Streptomyces lavendulae plasmid pSLG33
C:Species: Streptomyces lavendulae
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 07-May-1999
C:Accession: S34850
R:Feleberg, J.; Petricek, M.; Tichy, P.
Nucleic Acids Res. 21, 3582, 1993
A:Title: Nucleotide sequence of the mini-plasmid pSLG33 from Streptomyces lavendulae-g
A:Reference number: S34850; MUID:93348001; PMID:8346038
A:Accession: S34850
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-256 <FEL>
A:Cross-references: EMBL:X69872
C:Genetics:
A:Genome: plasmid pSLG33

Query Match 16.0%; Score 321.5; DB 2; Length 256;
Best Local Similarity 33.2%; Pred. No. 8e-19;
Matches 85; Conservative 45; Mismatches 105; Indels 21; Gaps 11;

QY 109 MVTMTHTAGORLHDLMTGLSAAWKAATNGRWR-----TERMYGCDGYVRAVEI 160
Db 1 MVTLTARHKGHHEPLFDVANGWRKLLSGAWAGDPKRGVLGERDLGVGNIRSLV 60
QY 161 THG-KNGWHVHVHALLMFSGDVSENILESFSDFMDR-WTSKLVSLGPAAPLRNSGGLDV 218
Db 61 TYGTRNGWPHLHVLLNNEETTEL--AYAMHRWDKTWRAWLKKGAFEPSEKH--GITW 116
QY 219 RKIGGEADQVLAAYLTATIASG-VGMEVSGDGKSGRHNAPWE-IAVDVGGDPQALEL 276
Db 117 SKV--TTPEEAGEVIAKAQKGVNTEARGDKRGLTGLAPPFEMLEYFRQTGMVVPV 174
QY 277 WREFPGSGRRATJAWSGRLAR-AGLGAELTAQIVEQESADVMVAIIPARSWMIRT 335
Db 175 WQEEKGTFRRAITWRSRLRAELGLDEELTDELAEEIGGETW-ALUPAESLRAIR 233
QY 336 CAPVVEIGLVFAG 351
Db 234 -VFGQSRILDTAENG 248

RESULT 2

A31844
rep protein - Streptomyces lividans plasmid pIJ101
C:Species: Streptomyces lividans
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:Accession: A31844
R:Kendall, K.J.; Cohen, S.N.
J. Bacteriol. 170, 4634-4651, 1988
A:Title: Complete nucleotide sequence of the Streptomyces lividans plasmid pIJ101 and c

A;Reference number: A91888; MUID:89008081; PMID:3170481
A;Accession: A31844
A;Molecule type: DNA
A;Residues: 1-456 <KEN>
A;Cross-references: UNIPROT:P22406
C;Genetics:
A;Genome: plasmid

Query Match 15.8%; Score 316; DB 2; Length 456;
Best Local Similarity 30.5%; Pred. No. 4.4e-18;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCCKGWCPCCKAGKVAHRADELSQVVAQL 102
DB 3 PASGVIVAQAAGTSVVL-----GLMRCGRILWCPVCAATIRHKRAEISITAAVVEVI 54
QY 103 GTGSAV-MVTMTMRTAGQRLHDLMTGLSAWKAATNGRR-----WRTE----- 145
DB 55 KRGGTAYLVTFARHGHTDRLADLMDALQGTTRKTPDPRPGAYQRLITGGTWAGRAKD 114
QY 146 -----REMYGCDGYRAVEITHGK-NGWHVHVHALLMFGSDV-----SENILSF 189
DB 115 GHRAADREGIRDIGVGMIRATEVTVGQINGWHPHIHAIVLVGKTEGERSAKQIVATF 174
QY 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGLDVRKIGGEAD-QVLAAYLTAKIA 237
DB 175 EPTGAALDEWQGHRSV-WTAALRKVNPAFTDDRHGVDKRLETERDANDLAETIAKTQ 233
QY 238 SG--VGMVEGSDGSGRGNAPWEI---AVDAVGG---DPOA-----LELWREPE 281
DB 234 DGKAPALEARADLKTATGNGVAPFELLGRLIGLGTGMDTDAAGVGSLENNLSRWHEYE 293
QY 282 FGSMDGRAIAWSRGLRARAGLGAELTDA 309
DB 294 RATRGRALIEWIRYLQMLGLDGDTEA 321

RESULT 3
S04020
hypothetical protein - Streptomyces sp. plasmid pSB24.2
C;Species: Streptomyces sp.
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: S04020
R;Bolotin, A.P.; Sorokin, A.V.; Alekandrov, N.N.; Danilenko, V.N.; Kozlov, Y.I.
Dokl. Biochem. 283, 260-263, 1985
A;Title: Nucleotide sequence of DNA of the actinomycete plasmid pSB24.2.
A;Reference number: S04020
A;Accession: S04020
A;Molecule type: DNA
A;Residues: 1-437 <BOL>
A;Cross-references: UNIPROT:Q52205; EMBL:X03756
A;Note: the authors translated the codon TGG for residue 59 as Arg, CGG for residue 187
A;Note: the source is designated as Streptomyces cyanogenus
C;Genetics:
A;Genome: plasmid

Query Match 11.5%; Score 230.5; DB 2; Length 437;
Best Local Similarity 24.6%; Pred. No. 4.1e-11;
Matches 102; Conservative 43; Mismatches 140; Indels 129; Gaps 19;
QY 55 GVTIVNGPKSGFG-GLRSCCKGWCPCCKAGKVAHRADELSQVVAQLGTGSA-MVTM 112
DB 6 GVTFARTAGAAVALGLLKCGRIWLCPCVCSQDPARFSEITEAVVSLMOQGGWAYLVTF 65
QY 113 TMRHTAGQRLHDLMTGL-----SAWKAATNGRRMTEREMY 149
DB 66 TARHTAADRLSDLMDALQGTADTETKRRPGAYQRLITGAAN--AGDKRKNQEGIR 123
QY 150 GCDGY---VRAVEITHGKN-NGWHVHVHALLMFGSDV-----SENILSFS---DAMP--- 194
DB 124 GRIGYIGMIRATEVTVGAGWHPHIHAIVLVGKTEGQRGDKRITGFTPTSEDALTWE 183

QY 195 DRWTSKLVSLGFAAPLRNSGL-----DVRKIGGEADQVLAAYLTAKIASGVG----- 241
DB 184 DRWERS-----GPATLARSTPGRPPTGARSPGATAGARATASTSSSVRSRQRPGRVHR 239
QY 242 -----MEVSGDGSGKGRHN-----RAPWEIAVDAVGGDPQ-----ALELW 277
DB 240 QDAGRQEPGAGTRRPGQRPQGHDLVLTETSRIG-DLMGCVPEEEAAGHGSGLAWGLDRW 298
QY 278 REPEFGSMG-----RRATAWSRGL-----RARAGL 302
DB 299 AEYETAIVSAGAHVRVDLPAPAGPDRRRHRGRHGRPVPDRRRFRDGVQWINDRAWNGL 358
QY 303 CAELTDAQIVQEESAPVWVAII-----PARSNMMIETCAPY-----VFEILG 346
DB 359 VGRSLDLAVVEAVEGREISMDALGELVQSAGSRAFLRLVLTPOEVTLEYDELLG 412

RESULT 4
I40335
hypothetical protein 1 - Corynebacterium glutamicum
C;Species: Corynebacterium glutamicum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40335
R;Yamaguchi, R.; Terabe, M.; Miwa, K.; Tsuchiya, M.; Takagi, H.; Morinaga, Y.; Nakamori
Agric. Biol. Chem. 50, 2771-2778, 1986
A;Title: Determination of the complete nucleotide sequence of Brevibacterium lactoferme
A;Reference number: I40334
A;Accession: I40335
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-240 <RES>
A;Cross-references: UNIPROT:Q45282; GB:D00038; NID:g216381; PIDN:BAA00026.1; PID:g216381

Query Match 8.4%; Score 169; DB 2; Length 240;
Best Local Similarity 32.0%; Pred. No. 2.2e-06;
Matches 54; Conservative 28; Mismatches 53; Indels 34; Gaps 10;
QY 193 MFDRTSKLVSLGFAAPLRNSG-GLD-VRKIGGEADQVLAAYLTAKIASGVMEVSGDGK 250
DB 1 MFSRWSAGVVGKAGMDAPLREHGKLVQDVSTWGGDAK-MATYLAK---GMSQELTGSATK 56
QY 251 SCRHNRAPEIAVDAVGGDPQALE-----LWREFEFGSMGRAIAWSRGLRARAGL 302
DB 57 TASKGSYTPFQM-LDMLAQSDAGEDMDAVLVARWEYEYVGSKNLKS-SWSRG--AKRAL 112
QY 303 GAEILTDAQI---VEQF-----ESAPVMAIIPARSMWIRT 335
DB 113 GDIYDADVRREMEELYKLGLEAPERVESTVAVALKPPDDWKLQIS 161

RESULT 5
B35390
replication protein - Lactobacillus hilgardii
C;Species: Lactobacillus hilgardii
C;Date: 31-Aug-1990 #sequence_revision 31-Aug-1990 #text_change 09-Jul-2004
C;Accession: B35390
R;Jossion, K.; Soetaert, P.; Michiels, F.; Joos, H.; Mahillon, J.
J. Bacteriol. 172, 3089-3099, 1990
A;Title: Lactobacillus hilgardii plasmid pLAB1000 consists of two functional cassettes
A;Reference number: A35390; MUID:90264300; PMID:2188951
A;Accession: B35390
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-314 <JOS>
A;Cross-references: UNIPROT:P35857; GB:M55222; NID:g149352; PIDN:AAA98163.1; PID:g149352
C;Superfamily: replication protein

Query Match 6.3%; Score 127; DB 2; Length 314;
Best Local Similarity 22.0%; Pred. No. 0.0081;
Matches 55; Conservative 50; Mismatches 107; Indels 38; Gaps 10;
QY 75 KGWTCPCCKAGKVAHRADELSQVY-AH-QLGTGSAVMTMTRHTAGQRLHDLWTGLSA 131

Db 69 KSLCLPLCNRRSMGOSNOLMQVLDEAHKORKTGRTFLTLTAENASGENLAKQEVKMG 128
QY 132 AKKAATNGRRWRTEREMYGCDGYVRAVEITHKNG-WHVHVALHMFSGDVSNILESPS 190
Db 129 A-----ISKLFQYKPAKALLGVRSSTEINKNGYHQHMHVLLFVKPTFKDSANYIN 183
QY 191 DAMFDRWTSKLSGLFAAPLNSGGDLVRKIGEGADQVLAAYLTAKIASGVGMEVSGDGK 250
Db 184 QAEWSKLQRAKMDY-QPVINVEARSNKAKGKSLIASAQET----- 226
QY 251 SRHGNGRAWEI-AVDAGDQQA-LELWREFFGSMGRRAIAMSRL--RAAGLAGAEL 306
Db 227 -----AKYQVSKDILNDQERDLQVVEDEFGLAGSRQISYG-GLFKEIRKQLQLED 278
QY 307 TDAQIVEOE 316
Db 279 VDAHLINVD 288

RESULT 6
S25721
replication protein repA - Streptococcus thermophilus
C:Species: Streptococcus thermophilus
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S25721
R:Janzen, T.; Kleinschmidt, J.; Neve, H.; Geis, A.
FEMS Microbiol. Lett. 95, 175-180, 1992
A:Title: Sequencing and characterization of pST1, a cryptic plasmid from Streptococcus
A:Reference number: S25721
A:Accession: S25721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <JAN>
A:Cross-references: UNIPROT:Q56129; EMBL:X65856; NID:G47880; PID:G47881
C:Superfamily: replication protein

Query Match 6.2%; Score 123.5; DB 2; Length 315;
Best Local Similarity 22.8%; Pred. No. 0.016;
Matches 50; Conservative 41; Mismatches 79; Indels 49; Gaps 11;

QY 12 KDRPVLVSSDKRGIRHELRLP-----KLOOI-----TTSETFNACGRPISGVNGVT 57
Db 5 KQGEILVDKNSRGKRDWRGRKILSLKLADIFKELQYKKTFFVERVISCGDTLQFIQND 64
QY 58 IVNGPKSGFGGLR-----SCGKGWICPCAGKVGAAHRADEISQV-----AHQLGTGSVM 109
Db 65 -----GNLKYQTYFC-KNKLCPMCNWRSMKYSYQTSRIVDEAIKQSPKGRFLF 113
QY 110 VTMTMRHTAGORLHDLWTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHKNG--W 167
Db 114 LTLTVKNVEGALNSTISQLTKSFDRLF--KRAKVQNRLL--GYLRSVEVTHNENKTY 168
QY 168 HVHVALHMFSGDVSNILESPSDAM-----FDRWTSKL 201
Db 169 HPHILVLM-----VRPSYFQSKDYITQKESDMSNSL 203

RESULT 7
T44800
replication protein repA [imported] - Streptococcus thermophilus plasmid pER35
C:Species: Streptococcus thermophilus
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44800
R:Solow, B.T.; Somkuti, G.A.
submitted to the EMBL Data Library, August 1999
A:Description: Comparison of low molecular weight heat stress proteins encoded on plasmid
A:Reference number: 222847
A:Accession: T44800
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-315 <SOL>
A:Cross-references: UNIPROT:Q9RNW4; EMBL:AFI77167; PID:AAF04355.1
A:Experimental source: strain ST135

C:Genetics:
A:Gene: repA
A:Genome: plasmid pER35
C:Superfamily: replication protein

Query Match 6.1%; Score 122.5; DB 2; Length 315;
Best Local Similarity 21.6%; Pred. No. 0.019;
Matches 48; Conservative 40; Mismatches 79; Indels 55; Gaps 9;

QY 12 KDRPVLVSSDKRGIRHELRLP-KLQOITTTSETFNACGRPISGVNGVTIYVNGPKSGFGGL 70
Db 5 KQGEILVDKNSRGKRDWRGRKILSLKLADIFKELQYKKTFFVERI----- 51
QY 71 RSCG-----KGMWICPCAGKVGAAHRADEISQV-----AHQLGTGS 106
Db 52 -SCGDTLQFIQNDGTLKYQAYFCNKKLCPMCNWRSMKYSYQTSRIVDEAIKQSPKGR 110
QY 107 VAMVTMRHTAGORLHDLWTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHKNG 166
Db 111 FLTLTVKNVEGALNSTISQLTKSFDRLF--KRAKVQNRLL--GYLRSVEVTHNEND 165
QY 167 --HVHVALHMFSGDVSNILESPSDAM-----FDRWTSKL 201
Db 166 KTYHPIHVLMM-----VRPSYFQSKDYITQKESDMSNSL 203

RESULT 8
JN0856
rep protein - Streptococcus thermophilus (strain No.29) plasmid pST1
C:Species: Streptococcus thermophilus
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: JN0856
R:Hashiba, H.; Takiguchi, R.; Joho, K.; Aoyama, K.; Hirota, T.
Biosci. Biotechnol. Biochem. 57, 1646-1649, 1993
A:Title: Identification of the replication region of Streptococcus thermophilus No.29
A:Reference number: JN0856; MUID:94080026; PMID:7764262
A:Accession: JN0856
A:Molecule type: DNA
A:Residues: 1-315 <HAS>
A:Cross-references: UNIPROT:Q9RNW8; UNIPROT:Q9RNW4
C:Genetics:
A:Genome: plasmid
C:Superfamily: replication protein
C:Keywords: replication initiation

Query Match 5.9%; Score 118.5; DB 2; Length 315;
Best Local Similarity 22.4%; Pred. No. 0.04;
Matches 52; Conservative 45; Mismatches 90; Indels 45; Gaps 12;

QY 12 KDRPVLVSSDKRGIRHELRLP-----KLOOI-----TTSETFNACGRPISGVNGVT 57
Db 5 KQGEILVDKNSRGKRDWRGRKILSLKLADIFKELQYKKTFFVERVISCGDTLQFIQND 64
QY 58 IVNGPKSGFGGLR-----SCGKGWICPCAGKVG--AHRADEISQVVAHQLGTGSVM 109
Db 65 -----GNLKYQAYFC-KNKLCPMCNWRSMKYSYQTSRIVDEAIKQSPKGRFLF 113
QY 110 VTMTMRHTAGORLHDLWTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHKNG--W 167
Db 114 LTLTVKNVEGALNSTISQLTKSFDRLF--KRAKVQNRLL--GYLRSVEVTHNENKTY 168
QY 168 HVHVALHMFSGDVSNILESPSDAMF--DRWT--SKL--VSLGFAAPLNSG 214
Db 169 HPHILVLMVRPSYFQSKDYITQAEWSDMWSLSKVDYVPMFIRTVKGTG 220

RESULT 9
A40651
replication protein A - Synechocystis sp. plasmid pCA2.4
C:Species: Synechocystis sp.
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: A40651
R:Yang, X.; McFadden, B.A.

J. Bacteriol. 175, 3981-3991, 1993
A;Title: A small plasmid, pCA2.4, from the cyanobacterium Synechocystis sp. strain PCC 6803
A;Reference number: A40651; MUID:93308076; PMID:8320214
A;Accession: A40651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <YAN>
A;Cross-references: UNIPROT:Q06460; GB:U13739; NID:g408891; PIDN:AAA02970.1; PID:g290013
C;Genetics:
A;Genome: plasmid
C;Superfamily: replication protein

Query Match 5.7%; Score 115; DB 2; Length 336;
Best Local Similarity 26.2%; Pred. No. 0.083;
Matches 34; Conservative 23; Mismatches 45; Indels 28; Gaps 6;

Qy 79 CPCGAGK--VGAHRADEISQVVAHQLTGSGVAMVTMRHTAGORLHDLWTGLSAANKA 135
Db 97 CPVCQWRSLMRKAKAFILQIVEAYPKHRFIFLTLVRNCELGQLRETTIGMNAWGR 156
Qy 136 ATNGRRWRTEREMGCDGVRAVEITHGKNG-WHVRVHALLM-----FS-GDVSENILES 188
Db 157 LVKRAKW-----PADGWIRSLVTRGKGSAAHFFHCLLMVKASYFSGHVISQEV--- 206
Qy 189 FSDAMFDRWT 198
Db 207 -----WT 208

RESULT 10
T46831
Plasmid replication protein [imported] - Streptococcus thermophilus plasmid pER341 in S. thermophilus
A;Species: Streptococcus thermophilus
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
A;Accession: T46831
R;Somkuti, G.A.; Solaiman, D.K.Y.; Steinberg, D.H.
A;Title: Structural and functional properties of the hsp16.4-bearing plasmid pER341 in S. thermophilus
A;Residues: 1-314 <SOM>
A;Reference number: Z24102; MUID:983252270; PMID:9657935
A;Accession: T46831
A;Status: preliminary; translated from GB/EMBL/DDDBJ
A;Molecule type: DNA
A;Residues: 1-314 <SOM>
A;Cross-references: UNIPROT:Q30850; EMBL:AF019139; NID:g3342702; PIDN:AAC27655.1; PID:g3342702
A;Experimental source: strain 134
C;Genetics:
A;Gene: rep
A;Genome: plasmid pER341
C;Superfamily: replication protein

Query Match 5.7%; Score 114; DB 2; Length 314;
Best Local Similarity 25.4%; Pred. No. 0.093;
Matches 47; Conservative 33; Mismatches 63; Indels 42; Gaps 10;

Qy 10 SGKDRPVLVSSDKGIRHLEKPLQOI-----TTSETFNACGEPISGVNGVTIVNG 61
Db 17 NGKDR-----DWRG-RKILSLKADIPKELQYKKTFFERVISCGDTLRFIQ----- 61
Qy 62 PKGSGGGGLR-----SGKGWICPCGAGKVGCAHRADEISQVW---AHQLGTGSGVAMVTMT 113
Db 62 ---KQDGSLLKYQAYFC-KNKLCPICNWRSMKSYQTSKIVDEAIKQEPKGRFLFTLT 117
Qy 114 MRHTAGORLHDLWTGLSAANKAAATNGRRWRTEREMGCDGVRAVEITHGK--NGWVHV 171
Db 118 VKNIEGKALNSTISQLTKSPDLRF--KRAKQVRNLL---GYLRSVEVTHNENDNSYPHI 172
Qy 172 HALIM 176
Db 173 HVLMM 177

RESULT 11
A49783

replication initiation protein rep [similarity] - Bacillus subtilis plasmids
N;Contains: DNA ligase [EC 6.5.1.1-]; DNA lyase [EC 4.2.99.-]
C;Species: Bacillus subtilis
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 09-Jul-2004
A;Accession: A49783; S70201
R;Hara, T.; Nagatomo, S.; Ogata, S.; Ueda, S.
Appl. Environ. Microbiol. 57, 1838-1841, 1991
A;Title: Molecular structure of the replication origin of a Bacillus subtilis (natto) pA49783
A;Reference number: A49783; MUID:91336725; PMID:1840479
A;Accession: A49783
A;Molecule type: DNA
A;Residues: 1-333 <HAR>
A;Cross-references: UNIPROT:Q00452; GB:I76715; NID:g151700; PIDN:AAA26037.1; PID:g151700
A;Experimental source: strain natto, plasmid pUH1
A;Note: authors translated the codon CCA for residue 45 as Ala, and GCA for residue 142
R;Meijer, W.J.J.; de Jong, A.; Bea, G.; Wisman, A.; Tjalsma, H.; Venema, G.; Bron, S.; Mol. Microbiol. 17, 621-631, 1995
A;Title: The endogenous Bacillus subtilis (natto) plasmids pTA1015 and pTA1040 contain
A;Reference number: S70199; MUID:96111483; PMID:8801417
A;Accession: S70201
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 294-333 <MEI>
A;Cross-references: EMBL:L26258; NID:g529115; PIDN:AAC37115.1; PID:g528977
A;Experimental source: strain IAM1028, plasmid pTA1015
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C;Genetics:
A;Gene: rep
A;Genome: plasmid
C;Superfamily: replication protein
C;Keywords: carbon-oxygen lyase; ligase; plasmid replication
F;245/Active site: Tyr (covalent DNA-binding) #status predicted

Query Match 5.6%; Score 111.5; DB 2; Length 333;
Best Local Similarity 25.5%; Pred. No. 0.16;
Matches 40; Conservative 29; Mismatches 61; Indels 27; Gaps 9;

Qy 78 ICPCGAGKVG---AHRADDEISQVVAHQLTGSGVAMVTMRHTAGORLHDLWTGLSAANK 134
Db 80 LCPMCANRSLKIAIYHKLIVVEANRQYCGWI-FLTLVRNVEGDLKPMIADMMKGM- 137
Qy 135 AATNGRRWRTEREMGCDGVRAVEIT--HGKNGWHVHVHALL-MPSGDVSENILESFSD 191
Db 138 ----NRLPAYKRVKVVATLGYFRALEITKNHEEDYHPHFVLLPVKSYFTTHYIKQ--- 190
Qy 192 AMFDRWTS---KLVSIGFAAPLNSGGLDVRKIGGEA 225
Db 191 ---SEWTSLWRAKMLDY-TPI-----VDIRRVKGRA 218

RESULT 12
A83561
Probable type II secretion system protein PA0683 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
A;Accession: A83561
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Li
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83561
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-381 <STO>
A;Cross-references: UNIPROT:Q915P2; GB:AE004503; GB:AE004091; NID:g9946553; PIDN:AAG04
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0683

Query Match 5.2%; Score 105; DB 2; Length 381;
Best Local Similarity 23.6%; Pred. No. 0.63;

Sat Oct 23 17:44:54 2004

A;Accession: AC3284
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-420 <KUR>
A;Cross-references: UNIPROT:O8YJ30; GB:AE008917; PIDN:AAL51438.1; PID:G17982146; GSPDB:G
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0256
A;Map position: I
C;Superfamily: Sarcosine oxidase
C;Keywords: oxidoreductase

Query Match 5.1%; Score 102.5; DB 2; Length 420;
Best Local Similarity 19.9%; Pred. No. 1.1;
Matches 63; Conservative 49; Mismatches 111; Indels 93; Gaps 12;
QY 95 SQVVAHQLGTGSVAVTMTMRTAGORLHDLWTGLSAAMKAATNGRRWR--TEREMYGCD 152
DB 152 AELAASQPGWDVREKAGIAYEHVGRRLAELQPLNPRFVAGTFVPGWKNVSDPKLF--- 208
QY 153 GYVRAVEITHGKNGWHVHVHALLMFSGDVSENILESFSDAMFDRWTSKLVSLGFAAPLNN 212
DB 209 -----GKAIW-----SYAESLGARFLSGKV---ASAKRQ 234
QY 213 SGGLDVRKIGGEADQVLAAYLTAKIAGVGMVEVSGDGKSGRHNRAPEIAVDVGGDPQ 272
DB 235 NGGVRLRLENG--EINATHVLVMAGAWSRDLAKGFG-----DIVELDT 277
QY 273 -----ALELWREFFEGSMGRRAIAWSRGLRARAGL--GAEI-----TDAQIVE 313
DB 278 RGYNTLPVGSFVVKQLTFPGHGFTTPTMETGLRVGGAVEFGGLDLPNFAFSEAMLKK 337
QY 314 QEESAPVMVAIIPARSMMIRTCA-----YVFG-ILGLVEAGATWEN 356
DB 338 ASKFLFGL-KVEGGRQWGYRPSMPDPLPVIGRASAGNVYGFHGHGLTQSAATARL 396
QY 357 LRDHLHYRLPAADVVP 372
DB 397 IRDLITGSEFAIDIEP 412

Search completed: October 23, 2004, 03:58:09
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 04:02:18 ; Search time 3866 Seconds
(without alignments)
3572.338 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGKDRPVLVS.....HLHYRLPAADVRRPIISVRK 379

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cn2_1/USPTO_spool/US10007527/runat_22102004_114116_26629/app_query.fasta_1.519
-DB=EST-QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10007527 @CGN 1.1 3437 @runat_22102004_114116_26629 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_hic:*
5: gb_est3:*
6: gb_est4:*
7: gb_est5:*
8: gb_est6:*
9: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	129.5	6.5	929	5	BUS19534
C 2	129	6.4	1648	4	BM913094
C 3	128.5	6.4	1090	4	BM457396
C 4	127	6.3	2033	9	AG347460
C 5	122.5	6.1	894	4	BL661991
C 6	122	6.1	2887	9	AV411933
C 7	120	6.0	918	9	CC623500
C 8	120	6.0	1255	9	CL479652
C 9	119.5	6.0	1661	8	CC321442

C 10	119	5.9	950	5	BU271357
C 11	119	5.9	1430	4	BM907785
C 12	119	5.9	1687	4	BI224374
C 13	118	5.9	851	5	BQ714786
C 14	118	5.9	893	8	BZ554247
C 15	117.5	5.9	1011	5	BX898881
C 16	117	5.8	1306	5	BQ936859
C 17	117	5.8	1615	2	BF783863
C 18	116.5	5.8	689	5	BU705875
C 19	116.5	5.8	695	7	CF730403
C 20	116.5	5.8	697	7	CF724491
C 21	116	5.8	623	7	CF433522
C 22	116	5.8	867	4	BI411296
C 23	116	5.8	943	6	CA280312
C 24	115.5	5.8	1199	4	BG684181
C 25	115.5	5.8	703	6	CB248284
C 26	115.5	5.8	786	5	BQ572893
C 27	115.5	5.8	799	6	CA320693
C 28	115.5	5.8	857	6	CA750314
C 29	115.5	5.8	918	6	BY717387
C 30	115.5	5.8	2161	3	AK049163
C 31	115.5	5.8	2161	3	AK083015
C 32	115.5	5.8	2162	3	AK028287
C 33	115.5	5.8	2167	3	BC013842
C 34	115.5	5.8	2168	3	AK017609
C 35	115	5.7	696	6	BY729572
C 36	115	5.7	899	8	BZ551248
C 37	114.5	5.7	745	9	CL829504
C 38	114.5	5.7	930	4	BG922360
C 39	114.5	5.7	1189	8	CC318464
C 40	114	5.7	787	7	CF878630
C 41	114	5.7	990	7	CF878672
C 42	114	5.7	1937	9	AG279610
C 43	113.5	5.7	974	5	BQ672428
C 44	113.5	5.7	1349	9	CL495597
C 45	113.5	5.7	1414	5	BQ276790

ALIGNMENTS

RESULT 1
BUS19534/c
LOCUS BUS19534 929 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10166703 NIH_MGC_134 Mus musculus cDNA clone
IMAGE: 6517370 5', mRNA sequence.
ACCESSION BUS19534
VERSION BUS19534.1 GI:22827060
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 929)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at:
<http://image.llnl.gov>
Plate: LHAM4097 row: g column: 03
High quality sequence start: 6
High quality sequence stop: 710.
Location/Qualifiers
1..929
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

/clone="IMAGE:6517370"
/tissue type="undifferentiated limb"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH MGC 134"
/note="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 0.0607 Length: 929
Score: 129.50 Matches: 91
Percent Similarity: 31.90% Conservativeness: 28
Best Local Similarity: 24.40% Mismatches: 95
Query Match: 6.46% Indels: 160
DB: 5 Gaps: 19

US-10-007-527A-2 (1-379) x BU519534 (1-929)

QY 20 SerSerAspLysArgGlyLeuArgHisGluLeuArgProLys-----Leu 34
DB 768 GCTTCTCAAAAGCGCTTGGCGCATCATCTT---CCTGAACACCCCTTGTGACTTTG 712
QY 35 GlnGlnIleThrSerGluThrPhe-AsnAlaCysGlyArgProLysSerGlyValAs 54
DB 711 CAAAAGGTACACAGTCTCAGAGCGATTGTGGAGCATGCTCACAA----- 669
QY 54 nGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyLeuArgSerCysG 74
DB 669 ----- 669
QY 74 yLysGlyTrpIleCysProCysAlaGlyLysValGlyAlaHisArgAlaAspGlu 94
DB 668 -----TGG---TGTCATGCTGC----- 654
QY 94 eSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThr 114
DB 653 -----GGAACATCTCACTAGTACCTGATCTGTAGGGT 619
QY 114 tArgHisThrAlaGlyGlnArgLeuHisAspLeuThrGlyLeuSerAlaAlaTrpLy 134
DB 618 A-----GGACAGGCTGGGAGACCA-----GGCTGAGCCAGACATGGAA 578
QY 134 sAlaAlaThrAsnGlyArgArgTrp-ArgThrGlu-----ArgGluMetTyGly-Cys 151
DB 577 G-----CGCTGTAGGAGGAGGCTTCTCGACACATGTTGCTGTGT 536
QY 152 AspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGly---TrpHisValHis 170
DB 535 -----GGGCACCTTGGCATCTGGGGTACTAGG 509
QY 171 ValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSer 190
DB 508 CTCACAGCATTTCTTGACTCGGC----- 482
QY 191 AspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeu 210
DB 481 -----TCTCAGGGCCCGTGCCTGGGCTCCCACTTCCGTGC 446
QY 211 ArgAsnSerGlyGlyLeuAspValArgLysIleGlyGlyGluAlaAspGlnValLeuAla 230
DB 445 CCGCCTCAGCAGACAGTGGAGCTCAGA-----GACGGAGTCTCTGA 404
QY 231 AlaTyrlleuThrLysIleAlaSerGlyVal---GlyMetGluValGlySerGlyAspGly 249
DB 403 GGGGATCTGCGCGGCTGTGACCAACATTAACAGTTGTACTCTTGTTCGTTGACGGG 344
QY 250 LysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaValGlyGly 269
DB 343 GGATGGGCAACCGGGGAGCCAGGCTCAGCTGGAGGGCTAGGGTCAGC----- 296
QY 270 AspProGlnAlaLeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArgAla 289

Db 295 -----TGG-----AGGCTGGG 284
QY 290 IleAlaTrpSerArgGlyLeuArgAlaAlaGlyLeuGlyAlaGluLeuThrAspAla 309
DB 283 CTCAGCTGGGGAACCTGGGCTCAGGAGAGACTGGGGCAGGA----- 242
QY 310 GlnIleValGluGlnGluSerAlaAlaProValMetValAlaIleIleProAlaArgSer 329
DB 241 -----GTCCAATCCAGCAGCCGTAACT-----CCGGCGCGC--- 209
QY 330 TrpMetMetIleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGlu 349
DB 208 -----GACGCGCGCGGCTGGTCACT 188
QY 350 AlaGlyAlaThrTrp-GluAsnLeuArgAspHisLeuHisTyArgLeuProAlaAlaAs 369
DB 187 CCGGGTTCGGCTGGGGCGCGGCTTGGGCTGTCTCCGGTGACACGACGAGGC 128
QY 369 pValArgProProlIleLeuSerValArgLys 379
DB 127 CACGAGGCCACACAGGCTCGGGGTGAGACAG 97

RESULT 2

BM913094/c
LOCUS
DEFINITION
5', mRNA sequence.
BM913094
VERSION
BM913094.1 GI:19363473
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1648)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM1989 row: g column: 07
High quality sequence stop: 263.

FEATURES
source
1..1648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5475102"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_41"
/note="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Alignment Scores:
Pred. No.: 0.151 Length: 1648
Score: 129.00 Matches: 93
Percent Similarity: 30.73% Conservativeness: 33

```
Best Local Similarity: 22.68% Mismatches: 153
Query Match: 6.43% Indels: 131
DB: 4 Gaps: 17

US-10-007-527A-2 (1-379) x BM913094 (1-1648)

QY 10 SerGlyLysAspArgProValLeuValSerSerLysArgGlyLeuArgHisGlu 29
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1101 GCTGCCACAGCGCCCGCACATCAATATCAC-----CCACTGGGGTCCGACACCA 1048

QY 30 LeuArgProLys-----33
Db 1047 AAGCGTCCAGCGGTGGCGGACGAGGAGCGTGTGACTGTCTGTCTTAGCAGATTG 988

QY 34 ---LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgPro-----49
Db 987 CCGTTGAGGCACTCGCCAGGNAG-----GGCGCGGTGAGACTGTA 946

QY 50 -----IleSerGlyValAsn-----54
Db 945 ATGTTGCTAGGGGGGTTAATACTCTGCGTGGCTGATGTATGTGACGAGGAGGACG 886

QY 55 -----GlyValThrIleValAsnGlyProLysGly 64
Db 885 GTGTTCAATTGTCGTCGCGCGCTTAGTAGTCGACAGAGGTGTGTGCGACCATGTGT 826

QY 65 SerGlyPheGlyGlyLeu-ArgSerCys-----GlyLysGlyT 77
Db 825 GCGGCTCTGTCTGGGCTACGTCGCTGCTGGGCTGGGAGTGCCTGCTGGGCGTGGAT 766

QY 77 rPileCysProCys-----CysAlaGlyLysValGlyAlaHisArgA 91
Db 765 GGGTGTGTTTGGCGCATTTGGAGGCGCGTGTGCGCGCATCTGTGCTGCTCGCGCG 706

QY 91 laAspGluIleSerGlnValValAlaHisGlnLeuGly---ThrGlySerValAlaMetV 110
Db 705 GATCTGTTTTTGGCTGCTCGCGGCACTAGCATGCGCGCGCGCGCTCCATCGGCA 646

QY 110 alThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrGlyLeus 130
Db 645 TTCAGCCGAGCGCGGCATCCGCC-----GAGGGGCGCG 610

QY 130 erAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgGluArgGluMetTyrG 150
Db 609 GGGGATCTGG---GTTGCGCGCCAGCGGCATGTGCGGCTGCGCGGCGACAGCTCG 553

QY 150 lyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGly-TrpHisVal 169
Db 552 GGGCGTGGGC-----GGTAGTCGCGGTTTGGCACTCC 520

QY 170 HisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPhe 189
Db 519 GAGACGGGGGATGGTGTGACTGTGCTGG-----490

QY 190 SerAspAlaMetPheAspArgTrpThrSerLysLeu-----201
Db 489 -----TGGACGGCGCTCAGCAGTTCGTGGCGGCTCGCCCGT 451

QY 202 ---ValSerLeuGly-PheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLy 220
Db 450 GGTGTGGCGGTGGCGTTTCCGAACGGCGCGCGCGTATAGTGGCAGCGTGGCGGCTTTC 391

QY 220 sIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVa 240
Db 390 TCGTGGCGGTTCGTTTGAAGTTCGCGGGGGTTCGTCACACAGTTTGTGCG-----336

QY 240 lGlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGly-----As 256
Db 335 -GACATGTGCGGGCGCGCGCGCGGTTCGCGGGAGCGCGGCTCCCGGACGAGAC 277

QY 256 nArgAlaProTrpGluIleAlaValAlaAspAlaValGlyAspProGlnAlaLeuGluLe 276
Db 276 GGAGGGGCGCTGG-----ACGCGGAGGGTGGCGCGCGCGCGGACCTGGCGCG 229
```

```
QY 276 uTrpArgGluPheGlu-PheGlySerMetMetGlyArgAlaIleAlaTrp-----292
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 228 GTGGCGCGCGCGCGCGACAGCGCGCGCGCTCGGCAAGCTGGGTCTCTGGAA 169

QY 293 -----SerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnI 311
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 168 GAGGCGCGAGAGCGCGCGGTGGGTGGGTCTGGAGCGGAGACCGCGAGGCG 109

QY 311 leValGluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSerTrpM 331
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 108 TCGCGCGCTCGGAGGCGCGGAGGCGGCTGTCAGCGGTGGCGCGCTCGCGCG 49

QY 331 etMetIleArgThrCysAla 337
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 48 CGGTAGCTCGATCCCGGCT 29

RESULT 3
BM457396 1090 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6407558 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5584861
DEFINITION 5', mRNA sequence.
ACCESSION BM457396
VERSION BM457396.1 GI:18506436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1090)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12350 row: d column: 14
High quality sequence stop: 538.
location/Qualifiers
1..1090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5584861"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores: 0.0949 Length: 1090
Pred. No.: 128.50 Matches: 80
Score: 31.38% Conservative: 27
Percent Similarity: 23.46% Mismatches: 94
Best Local Similarity: 6.41% Indels: 140
Query Match: 4 Gaps: 15
DB: 4

US-10-007-527A-2 (1-379) x BM457396 (1-1090)

QY 7 GluHisLeu-----SerGlyLysAspArgProValLeu-----18
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 314 GAGCACCTGCGAGCTAGCGCCAGGAGTCCCGCGGTCTAAAGAGCTGTGCAGATT 373
```

```
QY 19 -----ValSerSerAspLys 23
Db 374 GTGGAGGAGTATGGAGTGGTGGATCTACCGCTCTTCAGGGGTCTCTCCCAACATC 433
QY 24 ArgGlyIleArgHisGluLeu-----ArgProLysLeuGlnGln-IleThrTh 39
Db 434 CAGAGCTTCGCAGGAGATTGAGTCAGACGCGGAGCCAGACATCGCTCGGATGTTCAC 493
QY 39 rSerGluThrPheAsnAla-----CysGlyArgProLysSerGlyValAsnGlyVa 56
Db 494 CTCCAGACATTCACCTGCTCTCTCCCTGTCGCAAGCCCTATTCA-----539
QY 56 lThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGl 76
Db 540 -----AAGAA 544
QY 76 yTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGl 96
Db 545 CTGCCGGATCCCTGC-----TCACCTACCGGGCTCTAT-----GACAAATTGCTGA 592
QY 96 nValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHis 116
Db 593 AGCTGTAGGAATGCAATTCGAACTGAGCGCTTGGTCAACATCTAGAGGTGCTTCGGGA 652
QY 116 sThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTrp-----133
Db 653 ACTCCCTGTGCGAAACTACAGAAACCTTGGGAGTCCCTGATGAGGGCACCTGGTAACCA 712
QY 134 -----LysAlaAlaThrAsnGlyArgArgTrpArg-----143
Db 713 TGGCCCTCCATTGGGGGCACCAAGCAAGGTGTCATGTCACGACCTAGGAAAGGG 772
QY 144 -----ThrGlu-----ArgGluMetTyrGlyCysAspGlyTyrVa 155
Db 773 GGGGGGTACAGAACCCAGCTAAAGGGGATAAAGAACAGAAAGGATCGGGGGTTCAA 832
QY 155 lArgAlaValGluIleThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLe 175
Db 833 AAAGGAAAGCGGGGCACACCGAGGAAAGAGGGGGC-----869
QY 175 uMetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAs 195
Db 869 -----869
QY 195 pArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGl 215
Db 870 -----GGTCTGGCC-----AAGAAAAAGGGGG 892
QY 215 yLeuAspValArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrly 235
Db 893 AAACAAAGCGCGGAAAGGGGGGAAAA-----920
QY 235 sIleAlaSerGlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGl 255
Db 921 -----GGGGCGGAAAAAGGGCGGAAAAAGGGCGGAAAAAGGAAAGGAGCGGG 970
QY 255 yAsnArgAlaProTrpGluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGl 275
Db 971 G-----GTCGCGGAAAAAACAAGGGGGGGGAAAAAAGGG-----GA 1009
QY 275 uLeuTrpArgGluPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGl 295
Db 1010 ACAAAAAGGGGGGCGAAACACAGGGCGAAGCGGAAAGGGCGCAACGGGGGGTAAAGGGG 1069
QY 295 Y 295
Db 1070 G 1070
RESULT 4
AG347460/c 2033 bp DNA linear GSS 02-JUN-2004
LOCUS AG347460
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-142P21.TJ, genomic survey
```

```
sequence.
AG347460
AG347460.1 GI:47920770
GSS.
Mus musculus molossinus
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
BAC end Sequences of Library MSMg01
2 (bases 1 to 2033)
Unpublished
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Shuhiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsubura Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACe3.6
Vector : EcoRI
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
1..2033
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-142P21.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Alignment Scores:
Pred. No.: 0.315 Length: 2033
Score: 127.00 Matches: 92
Percent Similarity: 30.95% Conservative: 38
Best Local Similarity: 21.90% Mismatches: 126
Query Match: 6.33% Indels: 166
DB: 9 Gaps: 18
US-10-007-527A-2 (1-379) x AG347460 (1-2033)
QY 10 SerGlyLysAspArgProValLeuValSerSerAspLysArgGlyTle---ArgHis 28
Db 1930 GCTGGCGGCGCACACCGCCGATCGTCTCGG-GCTGTCGCGCTCGGCGTGTGTGCAT 1872
QY 29 GluLeu-----ArgProLysLeuGlnGlnIleThrThrSerGluThrPheAsnAla 45
Db 1871 CGGGTCTGTGGCGCGCGCCGAGGGCTT-----TACGGCGCT 1836
QY 46 CysGlyArgProLysSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65
Db 1835 GTCTGTCGGCGCTGTGTGGTTGCGGGCGGTGACGTGTGTGGCGGGGTGGCG 1776
QY 66 GlyPheGlyGlyLeuArg-----SerCys 73
Db 1775 GGTGACGTGGGGG-CGGGCGTCCCGCGTTCGTCGTCCTTCAGCGGTGTGTCTGTCG 1717
QY 74 GlyLysGlyTrpIle-----CysProCysCysAlaGlyLysVal 86
```



```
Db 1716 GGAGTCGGGTGTTGAGTCAGGTGGTGGCTGCTTGGTGTGCGCGCTGCGCGGTGT 1657
QY 87 GlyAla----- 88
Db 1656 GCGGCGCGCTCGAGCGCGCTTGGCGGGGCTGGGCGCGTGGCGCGCGGTGTGCGG 1597
QY 89 -----HisArgAlaAspGluIleSerGlnValAlaHisGln 101
Db 1596 GCGGCGTGGTGGCGCGGTACAGCCCTAGGCGATGCGGTGGGTGGCGCGGT 1537
QY 102 LeuGlyThrGly----- 105
Db 1536 CTGGGTGCGCGCTGCGTGTGGTTCGGGTGCTGCTCAGGTGGTATT 1477
QY 106 -----SerValAlaMetValThrMetThrMetArg 115
Db 1476 GGATTCGCGCGTGGTGTGAGCGTGGCGGAGTGGTGTGGCGCTGCGCGCGAGA 1417
QY 116 HisThrAlaGlyGlnArgLeuHisAspLeuTyr----- 126
Db 1416 GCGTCGCGCGGAGCTGTATCGCGCGGTGGGTGGTGGTGGCGCGGTGGTGGG 1357
QY 127 -----ThrGlyLeuSerAlaAlaTyrLysAlaAlaThrAsnGlyArg 140
Db 1356 CTTGAGGCTGCGGGGGTGGCGGTGGCGGTGGATGGATGCGCGCGCGGTTC 1297
QY 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLe 160
Db 1296 GCGTCTGCTAGCCCGG-----GTGAGACTG 1270
QY 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMet----- 176
Db 1269 CTGGGGGGGTGTGGATGGAGGTCGCGTGGCGGTGGCGCGGTGGCGCGAGCGG 1210
QY 177 -----PheSerGlyAsp----- 180
Db 1209 GCGGGGGTGGCGGAGGTGGTGGTGGGAGGCTTCGTCGGCGCGCGAGCGGGTG 1150
QY 181 ---ValSerGluAsnIleLeuGluSerPheSerAlaMetPheAspArgTyrThrSer 199
Db 1149 GCGGTAGCGCGGGGTGGT-----TTGTGGAGCGGG 1117
QY 200 LysLeuValSerLeu-----GlyPheAlaAlaProLeuArgAsnSerGlyGlyLeu 216
Db 1116 CGGCTCGTCGCGTGGGGGTGGCGGTGGCGGTGGTGGTGGCGCGCGCGGTGG 1057
QY 217 AspValArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIle 236
Db 1056 -----GCGGGGGGGGTGGCGCGGTGGCGCGGTGGCGCGCGCGCGCGGGCG 1009
QY 237 AlaSerGlyValGlyMetGlu-----ValGlySerGlyAspGly 249
Db 1008 TCGGGGGCGTGGAGTGTCTCCATCGCGCGTGGCGCGGTGGTGGTGGGTTC 949
QY 250 LysSerGlyArgHisGlyAsnArgAlaProTyrGluIleAlaValAspAlaValGlyGly 269
Db 948 GCTGGGGGTGGGGGGGTGCTCGCGCGGTGGTGGCGCGGTGGCGCGGTGGCG 889
QY 270 AspProGlnAlaLeuGluLeuTyrArgGluPheGluPheGlySerMetGlyArgArgAla 289
Db 888 GGGCCC---GCATGGCGGTGG-----GGTGGCGCGCTTCGCGGCC 847
QY 290 IleAlaTyr-----SerArgGlyLeuArgAlaArgAlaGlyLeuGly 303
Db 846 GGGCGGTGGCGGGGTGGGGCGCTCTGGCCCGCGGTGTGGCGCTCCCGGGGTGGGT 787
```

RESULT 5

```
BI661991/c
LOCUS 603304986F1 NCI_COAP_Mam4 894 bp mRNA linear EST 12-SEP-2001
DEFINITION 603304986F1 NCI_COAP_Mam4 Mus musculus cDNA clone IMAGE:5350797 5',
mRNA sequence.
ACCESSION BI661991
VERSION BI661991.1 GI:15576227
```

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 894)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM1891 row: k column: 22

High quality sequence stop: 849.

Location/Qualifiers

FEATURES

source

1..894

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NMRI"

/db_xref="taxon:10090"

/clone="IMAGE:5350797"

/tissue_type="tumor, gross tissue"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI_COAP_Mam4"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

ORIGIN

Alignment Scores:

Pred. No.:	0..276	Length:	894
Score:	122.50	Matches:	79
Percent Similarity:	30.77%	Conservative:	25
Best Local Similarity:	23.37%	Mismatches:	114
Query Match:	6.11%	Indels:	120
DB:	4	Gaps:	16

US-10-007-527A-2 (1-379) x BI661991 (1-894)

QY	68	GlyGlyLeuArgSerCysGlyLysGly-----Tyr-----IleCysPro	80
Db	811	GGGACATTGCACTCTCGCGGAAGCGATACGAGCTGGGATACACGCTTTTGATCTTGCCCA	752
QY	81	CysCysAla-----GlyLysValGlyAlaHisArgAlaAspGluLeuSer	95
Db	751	CATTGCGCTCTCTAAAGCGCTTGGCAATCTCTCTGGACACCTTGTCTGAC-----	701
QY	96	GlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArg	115
Db	700	-----TTTGGCAAGGTCTACAGTCTCAGAGCGATTGTGGAGCATGCT	659
QY	116	HisThr-----AlaGlyGlnArgLeu-----HisAspLeuTyrThrGlyLeuSer	130
Db	658	CACAATGTTGTCATGCTGCGGAACATCTCACTAGCAGCTGATCTTGTAGGTAGGAC	599
QY	131	AlaAlaTyrLysAla-----AlaThrAsnGlyArgArgTyrArg	143
Db	598	AAGGCTCTGGGAGCCAGGCTGAGCGAGCATGGAAGCGCTGGTAGGAGGAGCTTTCTC	539
QY	144	ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGly	163
Db	538	GACATCGTTGCTCTGTGGGCACCTTGGCATC-----	506

Qy 274 uGluLeuTrpArgGluPheGluPheGlySerMet--GlyArgArgAlaAlaAlaTrp--- 292
 Db 810 -CCTCTCTGGGAAGAAGATGTGCGCTCAGCCTCCGGGGCGAAGGTCTCTGAGGCTGGTG 752
 Qy 293 SerArgGlyLeuArg-AlaArgAlaGly----- 301
 Db 751 AGTGCTGCTCGCGAGGAGGTCTCGGGAGCGAGCCCGAGGTGGCGCTCTGGGGGTAGC 692
 Qy 302 -----LeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaPr 319
 Db 691 CATTGCCCTGGCGCGAGGAAGCTGGGCTCCCGATCAGCGACCTTGATGAGCATGCGCT 632
 Qy 319 oValMetValAlaIlelleProAlaArgSerTrp-----MetMetIleAr 334
 Db 631 CTTTGGCCCGAGGTTCACCGCAGCGAGGAGGTCTCTAGTAGAAGAGCTGGCGCTCG 572
 Qy 334 gThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTr 354
 Db 571 CACTGTAATCCGGTA-----AATCTCATAGTCTCTTCAGGGAACACTCTCTGGTG 521
 Qy 354 p 354
 Db 520 G 520

RESULT 7
 CC623500/c
 LOCUS
 DEFINITION OGV6046TM ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0501G19,
 genomic survey sequence.

ACCESSION CC623500
 VERSION CC623500.1 GI:31990145

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 918)
 AUTHORS Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGV6046TC
 Contact: Cathy Whitelaw
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

Location/Qualifiers

1..918

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0501G19"

/clone_lib="ZM 0.7 1.5 KB"

/note="Vector: pBSCK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 0.501 Length: 918

Score: 120.00 Matches: 66

Percent Similarity: 34.21% Conservative: 25

Best Local Similarity: 24.81% Mismatches: 89

Query Match: 5.99% Indels: 86

DB: 9 Gaps: 11

US-10-007-527A-2 (1-379) x CC623500 (1-918)

Qy 166 GlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIle 185

Db 796 GCGCGCAGCTCGCGCGCTACAGGTCTGTGGTTCAACGGCTGTACAGTCTCGACATC 737

Qy 186 LeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeu--- 204

Db 736 CTCGCCGGATGGACGACGCGTGGCGACGGCTCGACGCTCGCTCTCTCTCGGT 677

Qy 205 GlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGlyGlu 224

Db 676 GGGTTCCTCCATCCGCTCTTCGAG----- 653

Qy 225 AlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluVal 244

Db 652 --GACAGCATCGCATCGGCAGCTTCGCGCCACGCGCGCTCGCTCTCTCTCGTGTGC 596

Qy 245 GlySerGlyAspGlyLysSerGlyArgHis-----GlyAsnArgAlaProTrpGluIle 262

Db 595 CCCGCTGGGAACAACGGCCGCGGAGTCCGTCGCCAACGAGGCGCGTGGTGTCTG 536

Qy 263 AlaValAspAlaVal----- 267

Db 535 ACTGTGCGCGCGCCACCATGACCGCGCTTCGCGGTACGTCCGCTCGGTGACGGA 476

Qy 268 -----Gly 268

Db 475 CGTGTCTGTACGGCGAGTCCATGTACCCCGGGGAAACCGGTTTGAAGAAAGGC 416

Qy 269 GlyAspProGlnAlaLeuGluLeu-----TrpArgGluPheGluPhe 282

Db 415 GCGAAGGACCTGGAGCTGGAGCTAGTACGCGCTGGTGGGACCCGGGAATCCGAGTAC 356

Qy 283 -----GlySerMetGlyArgArgAlaIleAla-----TrpSerArg 294

Db 355 TGCCTCAAGGGTCTCTTGATAAGCCCGCTGCGCGGAAAGATGTTGGTATCGACCGC 296

Qy 295 GlyLeuArgAlaArgAla-----GlyLeuGlyAla 304

Db 295 GGCATCGCGCGCGCGCACAAAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 236

Qy 305 GluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProValMetVal----- 322

Db 235 GTGCTCACCAACTCCAGATTAACCGCGAGGAGACTCCGTCGACGCTCCAGTTCTGCCA 176

Qy 323 ---AlaIleIleProAlaArgSerTrpMetMetIle----- 333

Db 175 GCGACCTCATAGGGTACCGTGAGGCGCTGAGAGAGTACATCAGCTCGACGCCA 116

Qy 334 ArgThrCysAlaProTyrValPheGlyGluLeu---LeuGlyLeuValGluAlaGlyAla 352

Db 115 CGGCGGTGGCGAGGATCGTATTCGGCGGACGCGGATCGGACGAGCGCGCGCGCGCG 56

Qy 353 ThrTrpGluAsnLeuArg 358

Db 55 GTGCGCGGAATCTCGGCGC 38

RESULT 8

LOCUS CL479652/c

DEFINITION SAIL_30_C09.v1 SAIL Collection Arabidopsis thaliana genomic clone

ACCESSION CL479652

VERSION CL479652.1 GI:45949206

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1255)

AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,

1255 bp DNA linear GSS 01-APR-2004

SAIL_30_C09.v1 SAIL Collection Arabidopsis thaliana genomic clone

SAIL_30_C09.v1, genomic survey sequence.

Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M., and Goff, S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Sessions A

Applied Trait Genetics

Syngenta Biotechnology Inc.

3054 Cornwalis Rd., Research Triangle Park, NC 27709, USA

Email: allen.sessions@syngenta.com

ABRC Stock Number CS801460; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).

Sequences represent a pool of amplified genomic regions and not single contiguous sequences.

Class: TDNA tagged.

Location/Qualifiers

1..1255

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/clone="SAIL 30 C09.v1"

/clone_lib="SAIL Collection"

/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN

Alignment Scores:

Pred. No.: 0.773 Length: 1255
Score: 120.00 Matches: 94
Percent Similarity: 30.00% Conservative: 26
Best Local Similarity: 23.50% Mismatches: 106
Query Match: 5.99% Indels: 176
DB: 9 Gaps: 21

US-10-007-527A-2 (1-379) x CL479652 (1-1255)

QY 47 GlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySerGly 66
DB 1154 GGGCGGGTGAATTCGGGGTGTC-----CTGGTGGGGGCTTTAGAGGCAGGGGG 1107
QY 67 PheGlyGlyLeuArgSerCysGlyLysGlyTrp-----IleCysProCys--- 81
DB 1106 TGGTGTAGCGTGGGGCGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1047
QY 82 CysAlaGly-----LysValGly----- 87
DB 1046 TGTGCGGTGGCGCGCGCGGTCTGGGGCGGGCGGTGGGCATGTAGGCGTAGGTGGGG 987
QY 88 -----AlaHisArgAlaAspGluIleSerGlnValValAlaHisGlnLeuGlyThr 104
DB 986 CGCGGNAACGTCGTGATNGTCGGGAACGTCGCCCGCATGCCGGGACGCCAGCA-GGTAGC 928
QY 105 GlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisasp 124
DB 927 GGGGGTGA-----AGGAGTCCCGTAGTTCGACTA----- 895
QY 125 LeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArg-----Trp 142
DB 894 -----GCTGTCTGTAAGTGTGGTGGTGTG-TGTGGGAGCATNGGGGGGAATTGCTGG 842
QY 143 ArgThrGluArgGluMetTyrglyCysaspGlyTyrrValArgAlaValGluIleThrHis 162
DB 841 CCGTGTGGCGATGG-----GGACAAGGTCCGGGTAGTCAGCGCGCAGGAGG 797
QY 163 GlyLysAsnGly-----TrpHisValHisValHisAlaLeuLeuMetPheSer 178
DB 796 GGGCAAGGTGGCAGTGGGACAGATGGCGTGTGACAGGGGTAGTGGTGTCTG----- 746
QY 179 GlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThr 198

DB 746 ----- 746
QY 199 SerLysLeuValSerLeuGlyPheAlaPro-----LeuArgAsnSerGlyGlyLeu 216
DB 745 GGGCGGTGCGCCCGTGGGTCCGGGAGCCCGTGGGGCGTCCAGGAGCGGGGT--- 689
QY 217 AspVal-ArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrrLeuThrLysI 236
DB 688 ---GTGTAGTGGTGGGGGGAGCGG----- 664
QY 236 eAlaSerGlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArg----- 253
DB 663 -GGGGGGGGCGGTGGGTAAAGGGGGGGTGGGGGGCGACGCGACTGTGTGGAGGGAGAGGG 605
QY 254 -----HisGlyAsnArgAlaPro----- 259
DB 604 CGTGGCGGGGGGAGCGGGAGCGGGGGGGGGGAGGGGCGAGCGGGGAGCGGGGAGGG 545
QY 260 -----TrpGluIleAlaVal----- 264
DB 544 GGGGGGGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 485
QY 265 -----AspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGl 281
DB 484 GGGGGGGGGGGTGGGGGGGGTGGTCTCTTGTGGTGGTGGTGGGGTGGGGG----- 433
QY 281 uPheGlySerMetGlyArgArgAlaIleAlaTrp-----SerAr 294
DB 432 ---GGGGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 377
QY 294 gGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluCl 314
DB 376 GGGCGGTGGGGCGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 349
QY 314 nGluGluSerAlaProValMetValAlaIleProAlaArgSerTrpMetMetIleAr 334
DB 348 -----CG 347
QY 334 gThrCysAlaProTyrrValPheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTr 354
DB 346 TTCGTGGCGG-----GGGTGGGGCGCGTTCG 320
QY 354 pGluAsnLeuArgAspHisLeuHisTyrrArgLeuProAlaAlaAspValArgProPro 373
DB 319 CGGGGGTCTGGGTGGGACCCCGCGCGCGTGGGGGGGGGGGGGGGGGGGGGGGG 262
RESULT 9
CC321442 1661 bp DNA linear GSS 14-MAY-2003
LOCUS TAM32-34D5_ECL1.1 TAM32 Gallus gallus genomic clone TAM32-34D5,
DEFINITION genomic survey sequence.
ACCESSION CC321442
VERSION CC321442.1 GI:30715500
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1661)
AUTHORS Krawitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 152000 Std Error: 0.00
Seq primer: EC1 TACGACTCACTATAGGGCG
Class: BAC ends
High quality sequence start: 74

High quality sequence stop: 124.

FEATURES

source

1. .1661 LOCATION/ QUANTITIES

1661

```

1. 2002
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/cloned="TAM32-34D5"
/sex="female"
/cell_line="UCD001, inbred"
/cloned_lib="TAM32"
/clone="Vector: pCEBAC1; S
TAM32 Female Chicken lib
ordering information: htt

```

ORIGIN

Alignment Scores:

Pred. No.:	1.28	Length:	1661
Score:	119.50	Matches:	77
Percent Similarity:	33.23%	Conservative:	31
Best Local Similarity:	23.6%	Mismatches:	118
Query Match:	5.96%	Indels:	99
DB:	8	Gaps:	14

US-10-007-527A-2 (1-379) x CC321442 (1-1667)

Qy	54	AsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArg-----	71
Db	1492	GATGGCGTGACGGAGTCGGTGTGGAAAGGGGAGTGGATCAGGAGCCAGGAGGTGTC	1433
Qy	72	-----Ser	72
Db	1432	ATGTGGAGGGAGGGAGAAACGAGGTGCGGTTGTGGAGGGAGTGATAATGCAGGGAGT	1373
Qy	73	CysGlyGlyGlyTrrIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp	92
Db	1372	GTGGGAGAGGTAATAGGTGT-----GTGAGGTCCGGGAT	1337
Qy	93	GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMet	112
Db	1336	ACGAGAGGTGTGTGGTGGTCATGATCGTAGGGCGAGGGGCGGGTCGGAGTGGGGCTCG	1277
Qy	113	ThrMetArgHisThrAlaGlyGlnArgLeu-HisAspLeuTrpThrGly-----	128
Db	1276	GGGGGTGGAGTGGTGAAGGGGACACGGGTAGAGG-----TGGGGAGGTGGGGTGGGAG	1223
Qy	129	----LeuSerAlaAlaTrpLys---AlaAlaThrAsnGlyArgArgTrpArgThrGluAr	146
Db	1222	GAGGTGGGTGCAGGTGGAGGGAGGTGGATGAGAGGAGGGCGGTGGAGTTGGAGGG	1163
Qy	146	gGluMetTyrgLysCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGl	166
Db	1162	GGAGGCGCGAGTCAGGATGTTGGCTGTAGCGGGTG-----GGCGG	1121
Qy	166	YTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLe	186
Db	1120	GTGG-----TTGTGCGTGGGAGCAGCGTGCAGCGGAGGGTGCAG	1082
Qy	186	uGluSerPheSerAsp-AlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyP	206
Db	1081	CAGGAGGGGGGAGAGAGATGGAGCGGAGGGGAGCAGCAGCAGCGGTGGGG--GGGA	1025
Qy	206	heAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIle--GlyGlyGluA	225
Db	1024	GGCGGGGGGGGAGTGTGACGAGAGGGGTGGAGGACGGGGGTGGGGCGGGGAGC	965
Qy	225	laAspGlnValLeuAlaAalatrLeuThrLysIleAlaSerGlyValGlyMet-----	242
Db	964	GGGAGCGAGGGTTGGCGGGTTGGTGTGGGG--GGGAGTGGGATGGTATGGTGGGTG	908
Qy	242	-----	242
Db	907	AGGGGGGGTCTGTGTAGGGGGGGCGGAGCGAGCGAGGGGAGTGGAGGGGGGGGA	848

Qy	243	-----GluValGlySerGlyAaspGlyLySerGlyArgHisGlyAasnArgA	258
Db	847	GTTCGTGGCAGGGGAGGTGGGAGGGGTCGGGGGGGAGGTGAGGGGGGGG	788
Qy	258	laProTTPGluileAlaValaaspAlaValGlyGlyAaspProGlnAlaLeuGluLeuTTP-	277
Db	787	GGGGGTGGAGAGGTGTGGGGGTGGGCAGGGGGGGCTGAGAGGGGGG	734
Qy	278	-----ArgGluPheGluPheGlys	284
Db	733	GGGGGTGGGGTAGTGGGCGGAGAGTTGCGGGAGCGGTGGGAGTCGGAGTAGGGGG	674
Qy	284	er----MetGlyArgArgAlaileAlaTTPSerArgGlyLeuArgAlaArgAlaGlyLeuG	303
Db	673	GGTGGGTGGGGCGAGGAGAGGGGGGGGTGGGGCAGGGGGGGCCAGCGGAGGGGGTGTGG	614
Qy	303	lyAlaGluLeu 306	
Db	613	GGGGCCGAGTG 603	
RESULT 10			
BUZ71357/c			
LOCUS	BUZ71357	950 bp	linear
DEFINITION	603818506F1 CSEQHN52 Gallus gallus	cdna	clone
	sequence.		EST 26-NOV-2000
ACCESSION	BUZ71357		
VERSION	BUZ71357.1	GI:25542307	
KEYWORDS	EST.		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;		
	Phasianinae; Gallus.		
	1 (bases 1 to 950)		
REFERENCE	Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.		
AUTHORS	Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.		
TITLE	A Comprehensive Collection of Chicken cDNAs		
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)		
MEDLINE	22335534		
PUBMED	12445392		
COMMENT	Contact: Simon Hubbard		
	Department of Biomolecular Sciences		
	University of Manchester Institute of Science and Technology		
	(UMIST)		
	PO Box 88, Manchester, M60 1QD, UK		
	Tel: 01612008930		
	Fax: 01612360409		
	Email: Simon.Hubbard@umist.ac.uk.		

FEATURES

source

```

1. 330
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, H
/db_xref="taxon:9031"
/clone="ChEST816x2"
/dev_stage="22"
/lab_host="DH108"

```

Vector: pBluescript II KS(+); Site1: NotI; Site2: NotI; this normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylation C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9248-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.: 0.657 Length: 950
Score: 119.00 Matches: 80
Percent Similarity: 30.08% Conservative: 31
Best Local Similarity: 21.68% Mismatches: 120
Query Match: 5.94% Indels: 139
DB: 5 Gaps: 17

US-10-007-527A-2 (1-379) x BU2711357 (1-950)

QY 43 PheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyPro 62
DB 909 TAAGTCTTTCCCGCGGTGTGTGGGGGTGTCCGGGGGTGCACGTGGCCCTCGGTTGC 850
QY 63 LysGlySerGlyPheGly-----GlyLeuArgSerCysGlyLysGlyTyr 77
DB 849 TTGGGGGGTGGGTTTGGTGTGTCTTTTGGGGGGCGGAGGGGTGCCATGGGGATGG 790
QY 78 IleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal 97
DB 789 GTC-----TGCTGTGGCTCGGGTGTGGG----- 766
QY 98 ValAlaHisGlnLeuGlyThrGlySerVal-AlaMetValThrMetThrMetArgHisTh 117
DB 765 -----CTTGGGATGGGCCCTTGTGTCTGCTGCGTGGTTGG----- 732
QY 117 rAlaGlyGlnArgLeuHisAspLeuThrGlyLeuSerAlaAlaTyr-LysAlaAlaT 137
DB 731 -----GTTTGGTTTGGT-----TGTTGGTTTGGCTGG 703
QY 137 hrAsnGly-----ArgArgTyrArgThrGluArgGluMetTyrGlyCysAspG 153
DB 702 CAGGGGGTTTTGTGTGCTTGTGGTGGGGCGTGCATGCGCTGGTGGCGCTGGGG 643
QY 153 lYtyrValArgAlaValGluIleThrHisGlyLysAsnGly-----T 167
DB 642 GCAGGGTTCGTGTGCTTCTTCTATCATGCGGGCGGGGGTCTTCTGGGCTGAGGTGT 583
QY 167 rPhisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuG 187
DB 582 GG-----GGGGAGGTTTGGCCAGGG----- 563
QY 187 luSerPheSerAspAlaMetPheAspArgTyrThrSerLysLeuValSerLeuGlyPheA 207
DB 562 -----TGTTGGGGGCTGG 550
QY 207 laAlaProLeuArgAsnSerGlyLeuAspValArgLysIleGlyGluAlaAspG 227
DB 549 TTTGAGGCTCGGCTGGGTTTGGTGGCTCCGGGTTTCGGCTGGTTCAGGTTGTGCTGGG 490
QY 227 lnVal-----LeuAlaAlaTyrLeuThrLysIleAlaSerGlyValG 241
DB 489 CCGTGAGTTTCAGTGGCTGGGGCCATCTCAGGTCAGGCTTAGTTGTGCGAGCCTTGGGATCAG 430
QY 241 lYMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrog 261
DB 429 GCGTGGAGCAGGGGTGTGA-----TGCC 406
QY 261 luIleAlaValAspAlaValAlGlyLysProGlnAlaLeuGluLeuTyrArgGluPheG 281
DB 405 GAGTGCAGGTAGGAGGCAAGTCTGGGCAA----- 377
QY 281 luPheGlySerMetGlyArgArgAlaIleAlaTyr-----SerArgGlyLeuArgAlaA 299
DB 376 -----GGCCCCGAGGCAAGGCCGTAGGCTTGGTATCGTTGCATGGGGATCGCAGCA 322
QY 299 rGAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaP 319
DB 321 GAGCA-----TCAGCGC 310
QY 319 roValMetValAlaIleIleProAlaArg-----SerTyrMetMetIleA 334

Db 309 CGTGGGCTCAGAGTTTGTCTTCAACGCGAGGTTTCAGATGAAGCTTCATGGCTTCAGGCAG 250
QY 334 rgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGlu----- 349
Db 249 CAAAGAAGCTGAGTTGTCAGTTGTCAGGTCAGTCTGTGGGGCCAGCAGAAACACAGGACA 190
QY 350 --AlaGlyAlaThrTyrTrpGluAsnLeuArgAspHisLeuHisTyrArg-----LeuP 366
Db 189 GGGCTGGAAACCATCAGATAGCATTCCTTTTCAGCCAGCATTTAAAGCCCTGCGCTGCC 130
QY 366 roAlaAlaAspValArgProPro 373
Db 129 CTGCAGCATG-GAGCGTCCGCCT 108

RESULT 11

BM907785/c

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

BM907785

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1430)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DP/Gapdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12/44 row: f column: 10

High quality sequence stop: 319.

FEATURES

source

1. .1430

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5736201"

/tissue_type="large cell carcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 68"

/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 1.8 kb. Library constructed by Life

Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 1.16 Length: 1430

Score: 119.00 Matches: 90

Percent Similarity: 34.82% Conservative: 43

Best Local Similarity: 23.56% Mismatches: 110

Query Match: 5.94% Indels: 139

DB: 4 Gaps: 21

US-10-007-527A-2 (1-379) x BM907785 (1-1430)

QY 46 CysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65

Db 1343 TGTGGGGGCCCA-----GGTATGGGGGGTATGG-----GGGGGTGGGGGAAA 1299

QY 66 GlyPheGlyLeuArgSerCysGlyLys----- 75

```
Db 1298 GGTGTGGGTGGTATTGTTTAATGTGTCCGTCCTTATGTAGGGGGTGGCGGGGG 1239
QY 76 -----GlyTrpIleCys 79
Db 1238 GTGGATATCAGAGTGGGGGTACGTGCCCTCATATGACAGCGGTGGGTAATTTTGT 1179
QY 80 ProCys-----CysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1178 CAGTGTGGCGAGGATATGGGTATATTACGGGGGGCGATTTCGGA-----AGGTGGAGN 1125
QY 93 GluIleSerGlnValValAlaHisGlnLeuGly-ThrGlySer-----ValAlaMetVa 110
Db 1124 GNGCTCTCTAGTCTGTATTGCGGTATATCGCCATAGGCGCTGATGAATAGAGCTCGT 1065
QY 110 lThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeu-- 129
Db 1064 GACTATCTCT---AAACAT-----TTGTGGGTGGTGTGTG 1032
QY 130 -----SerAlaAlaTrp-----Lys-AlaAlaThrAsnG 139
Db 1031 CGGGACCGGGCGAGTCAGTGGCGCTTTAGTGGGTTTGTGCGCGTGGGATAACGTCTG 972
QY 139 lYArgArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValG 159
Db 971 GTAGCCGCGCGCGGTGAGAGAGCGAGTTTAAATGTGACGG-----GCTTTAT 921
QY 159 luIleThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuMetPheSerG 179
Db 920 GTCTAGACTAGGCGCTTAGAGGGGTGGCGAGGGTGTGTACGAGCTTTTGTGTGTGG 861
QY 179 lYAspVal-----SerGluAsnIleLeuGluSerPheS 190
Db 860 GGGCGGTAGGATAATGTGCGCGGTAAAGAGAGTCCAGAAATGGTGGATTCGGCATAT 801
QY 190 exrAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaProL 210
Db 800 CGAACAGTGAT-----GTCAGCCCTATATATGCG-----G 771
QY 210 euArgAsnSerGlyGlyLeu-----AspValA 219
Db 770 TAGCGGAATCAGCGCGCTGTAATGAGTCCATAAGGGTGTCCGANAAGTCCATGAGTGT 711
QY 219 rgLysIleGlyGly-----GluAlaAspGlnValLeuAlaA 231
Db 710 GCCTGATAGTGTATTTTCCAGCAATATGTGAGGTGGAGGTGATAAGAGATGGGGA 651
QY 231 laTyrLeuThrLysIleAlaSerGlyValGlyMet----- 242
Db 650 GTATAGGTGTGTAGTGGAAATCAGGGGAAGGAGTTTGGGGTAGCGAGTGACGTAATGG 591
QY 243 -----GluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpG 261
Db 590 CGAGGGCGCAGGGTTTGGCGGAGGGGATAGGGGTGATGTGGGTGGCAGCGCTCCCTTT 531
QY 261 luIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheG 281
Db 530 GCCTATGGTGGCGAGCGTGTGT-----GTTGAGTGGCTCAGGAGATTG 486
QY 281 luPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGly-----LeuA 297
Db 485 GGGGAGGTAGTGTGGATGGCTGTATTGGGGCAGTAAAGAGGGGATGATGTTGTGGGC 426
QY 297 rg-AlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGlu 316
Db 425 GTGGCGAGGGGGGCGAGCGGC-----GCGGGGTGTGTGGCGAGGT--- 381
QY 317 SerAlaProValMetValAlaIlelleProAlaArgSerTrpMetMetIleArgThrCys 336
Db 380 -----CCGGGCGGTGGTGGTGTACTAGAGTGTGT 348
```

RESULT 12
BI224374/c
LOCUS

1687 bp mRNA linear EST 11-JUL-2001

```
DEFINITION 602940916Fl NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104346 5',
mRNA sequence.
ACCESSION BI224374
VERSION BI224374.1 GI:14677818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1687)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E.B. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11251 row: 0 column: 03
High quality sequence stop: 195.
FEATURES
source
1..1687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5104346"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/notice="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
```

ORIGIN

Alignment Scores:

Pred. No.:	1.46	Length:	1687
Score:	119.00	Matches:	81
Percent Similarity:	33.72%	Conservative:	34
Best Local Similarity:	23.75%	Mismatches:	102
Query Match:	5.94%	Indels:	124
DB:	4	Gaps:	18

US-10-007-527A-2 (1-379) x BI224374 (1-1687)

```
QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGly-----Trp 77
Db 1036 GGGAGGGTGGGAGTGGAGCGAGTGGG-----GGCTGGCGAGAGGCTGTGCGAGGTGG 983
QY 78 lIleCys-----ProCysCysAlaGlyLys-----ValGlyAlaHisArg 90
Db 982 CTGTGTGATGACGGCGGCAGTGTGTGCGCGAGGGCGCGCGGAGTGGGTGCGCCGC 923
QY 91 AlaAspGluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetVal 110
Db 922 GCGGCAGAGGTG---AAGGGAGGGCGAGCCGACCGAGGGGTAGAGGAGCG----- 878
QY 111 ThrMetThrMetArgHisThrAlaGlyGlnArg---LeuHisAspLeuTrpThrGlyLeu 129
Db 877 -----GCTGGAGCACGCTGAGGAGAGTGTGTGTGGGGGGGGCTG 839
QY 130 SerAlaAla-TripLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTy 149
Db 838 ACATGAGCAGTGGCGAGGGCGGAGAGCGGTGCCCTGGAGGAGC----- 793
QY 149 rgCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGlyTrpHisVa 169
Db 792 -----AGGGCGGTGGCGA--- 781
```



```

QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 780 -----CATGCT----- 775
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 774 -----GGCAACAGCGGCC 761
QY 209 oLeuArgAnSerGlyGly----- 215
Db 760 GCGCAGCAACCGCGGAGAGCGCGCGGAGAGGAGAGAGAGAGAGAGAGAGTGG 701
QY 216 -----LeuAspValArgLysIleGlyGlu-----AlaAspGlnValLeuAl 230
Db 700 GGAGCCAGTGGGAGCGCGCGCGAGAGTGGGTGACAAAGTGGCGGC 641
QY 230 aLaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySer----- 246
Db 640 G-----GAGAGCGCGCGCGGTGTGAGCAGGAGAGCCCGTGGGAGAA 599
QY 247 ----GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAs 265
Db 598 GATGGTGGGGCGAAATGAATAGACACAGCGAGCGAGCTGCGCGAGCGGAGGTG 539
QY 265 pAlaVal-----GlyGlyAspProGlnAlaLeuGluLeuTipArgGluPh 280
Db 538 GGTGACGAGAGGCTGTCGAGAGGGGGCAACGGCGGAA-----TGGAGCGCGG 488
QY 280 eGluPheGlySerMetGlyArgArgAlaIleAlaTrp-----SerArgGlyLeuArgAl 298
Db 487 GCGCGCGGGGCTGAGGAAAGGGGGGAGGATGCTGTGACGCGAGCAATGTTGTGCGGC 428
QY 298 aArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerA 318
Db 427 AGCTAGGGCCATGCGCGGAGATGCGAGCGCGGAGCAGCGAGCGGCGGAGAG 368
QY 318 laProValMetValAlaIleIle-----ProAlaArgSerTrpMetIleArgT 335
Db 367 CGCGCTGGGGAGCGAGGCGAGGCGTGGAGCTGCGAGCGCGAGCGCGAGTGGTGGCGGAG 308
QY 335 hrCysAlaProTyrValPheGlyGlyLeuValGluAlaGlyAlaThrTrp 354
Db 307 TGTGACGCGCAGCGGTGCCAGCAGGTGTGTGTGTTCCATGGGGAATTAGAGTGTGG 249

RESULT 13
BQ714786/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
COMMENT

BQ714786 851 bp mRNA linear EST 16-JUL-2002
AGENCY: 8482181 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6307476
5', mRNA sequence.
BQ714786
BQ714786.1 GI:21853685
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 851)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: ResGen, Invitrogen Corp
DNA Sequencing by: the I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13725 row: a column: 13
High quality sequence stop: 666.
Location/Qualifiers
1. .851
FEATURES
source

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6307476"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 129"
/notes="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 2.2 kb. Constructed
by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
Library."

ORIGIN
Alignment Scores:
Pred. No.: 0.705 Length: 851
Score: 118.00 Matches: 78
Percent Similarity: 32.43% Conservative: 30
Best Local Similarity: 23.42% Mismatches: 112
Query Match: 5.89% Indels: 113
DB: 5 Gaps: 15

US-10-007-527A-2 (1-379) x BQ714786 (1-851)
QY 68 GlyGlyLeuArgSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGly 87
Db 845 GCGGGAACCGATACGACCGTGGGGATAC-----ACCGTTGATTG 801
QY 88 AlaHisArgAlaAspGluIleSerGlnValAlaHisGlnLeu----- 102
Db 800 GCCACATGGCG-----CTCTCAAAGCGTTCGCGCATCATCTCTGAACCCCTTGCTT 747
QY 103 -----GlyThrGlySerValAlaMetValThrMetThrMetArgHisThr----- 117
Db 746 GACTTTGGCAAGGTCACAGTCTCAGAGCGATGTTGGAGCATGCTCACAATGGTGTCCAT 687
QY 118 AlaGlyGlnArgLeu-----HisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAla 135
Db 686 GTTGGGAACATCTCAACTAGCACCTGATCTTAGGTAGGACAGACCTGGGAGACC 627
QY 136 -----AlaThrAsnGlyArgArgTrpArgThrGluArgGluMet 148
Db 626 AGGCTGAGCAGCAGCATGGAAGCGCTGGTAGCAGGAGCTTCTCGACACATGGTTGCTC 567
QY 149 TyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGlyTrpHis 168
Db 566 TGTGGCACCTTGGCATC-----TGGGT 543
QY 169 ValHisValHisAlaLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSer 188
Db 542 ACTAGGCTCCACAGCATCTCTTGGACTCGGC----- 510
QY 189 PheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAla 208
Db 509 -----TCTAGGCGCGCTGCTGGCTCCCAACTT 480
QY 209 ProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGlyGluAlaAspGlnVal 228
Db 479 CCGTCCCGCTCAGCAGGAGTGGAGCTCAGA-----GACGAGATC 438
QY 229 LeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal---GlyMetGluValGlySerGly 247
Db 437 CTCTGAGGGGATCTTCCCGGCTGTGACCAACATAACAGTTGTACTCTTTGTCGCTT 378
QY 248 AspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVal 267
Db 377 GACGCGGATGGGCAACGCGGCGGCTCAGCTGAGGGCTAGGCTCAGC----- 324
QY 268 GlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArg 287
Db 323 -----TGG-----AGG 318
QY 288 ArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaGlyLeuGluLeuThr 307
Db -----

```

```

Db 317 GCTGGGCTCAGCTGGGGAAGCTGGGCTCAGGCAGAGAACTGGGGCAGGA----- 270
Qy 308 AspAlaGlnIleValGluGlnGluSerAlaProValMetValAlaIleProAla 327
Db 269 -----GTCAATCCAGGAGCGCGTAACCT-----CCGGCG 240
Qy 328 ArgSerTrpMetMetIleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeu 347
Db 239 CGC-----GAGCGCGCGGGCTG 222
Qy 348 ValGluAlaGlyAlaThrTrp-GluAsnLeuArgAspHisLeuHisTyrArgLeuProAl 367
Db 221 GTCACTCCGGGTTCCGGCTGGGGCCGGCGGCTTCCGGCTGTGCTCCGGGTGAAGCG 162
Qy 367 aAlaAspValArgProProIleIleSerValArglys 379
Db 161 AGGAGCCAGGAGCCACCAGGCGCTCGGGGTGAGACAG 125

RESULT 14
BZ554247/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 893)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
location/Qualifiers
1..893
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacs1-60_4590"
/clone_lib="pacs1-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."
ORIGIN
Alignment Scores:
Pred. No.: 0.754 Length: 893
Score: 118.00 Matches: 74
Percent Similarity: 34.07% Conservative: 19
Best Local Similarity: 27.11% Mismatches: 116
Query Match: 5.89% Indels: 66
DB: 8 Gaps: 10

US-10-007-527A-2 (1-379) x BZ554247 (1-893)

Qy 16 ProValLeuValSerSerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGln 35
Db 730 CCAGTCTGTCCCGCGGAACCCCTTCGGGGCGTTCAATTCGGGAC-AGCCAGAGCCGACG 672
Qy 36 GlnIleThrThrSer-----GluThrPheAsnAlaCysGlyArgProIle 50
Db 671 CAGCGACCCCGCAGCATGCTCGCCAGTGCTCGGCTTTTCAGCGCGGTGAACGGCGCGG 612

```

```

Qy 51 SerGlyValAsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeu 70
Db 611 GTTGGC-----GTAGNCCCGAGCAGCGCGCGCGCGCGAGCGGGTC 570
Qy 71 ArgSerCysGlyLysGlyTyrIleCysProCysCysAlaGlyLysValGlyAlaHisArg 90
Db 569 CAGNCT-----GCCGTCGCGATTGGAGCACCAGTC 540
Qy 91 AlaAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetVal 110
Db 539 GTTCAGGGTGATCCAGGTCGTAGAGCATCCAGCGGAGCAGGC---GTTGTAGAAGTC 483
Qy 111 ThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTyrThrGlyLeuSer 130
Db 482 GATCAGACCGCCAGGTGCGGCGCTCGAACACAGCA-CGTTGTCGGGAACAGGTCCGGCAT 424
Qy 131 AlaAlaTyrLysAlaAlaThrAsnGlyArgArgTyrArgThrGluArgGluMetTyrGly 150
Db 423 GCAGTTGGCGCGCGGAGCGCGCAGCGGCTCGGCATCCAGCGCGCGGATCTCCGCCAGGG 364
Qy 151 CysAspGlyTyrValArgAla---ValGluIleThrHisGlyLysAsnGlyTyrHisVal 169
Db 363 CCGGGCGGAGCAGCGCGGCGGCTGTTCCGGTAGTCGCGGGCGCA---GGTTGGCAC--- 310
Qy 170 HisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPhe 189
Db 309 -----CTTGTTCACGA 298
Qy 190 SerAspAlaMetPheAspArgTyrThrSerLysLeuValSerLeuGlyPheAla----- 207
Db 237 TCCAGGGCAGACCCAGGTCGCTGGGGCTTCGAGGATCGCGCGGGTGGCGGTGAA 238
Qy 208 -----AlaProLeuArgAsnSerGlyLeuAspValArgLysIleGly 222
Db 237 GGTGGCCGAGCAGATCGCCACTTCCTGGCAGTGTGGCGCTCGGCTGCGCGCTGTGGC 178
Qy 223 GlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMet 242
Db 177 GGCCCGCCAGCGCGGTTGCAGCAGCG-----CCGGCTTGCCTT 139
Qy 243 GluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrGluIle 262
Db 138 CGAGACGGCGAGCGCCTCGCGCTCGCGGTGCGGAGCAGCGTAGGCGCACCG----- 88
Qy 263 AlaValAspAlaValGlyGlyAspProGlnAlaLeuGlu 275
Db 87 -----GGGATCCACTAGTTCTAGAG 67

RESULT 15
BX898881/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 1011)
Han, B., Feng, Q., Huang, Y.C., Ying, K., Li, Y., Guan, J.P., Zhu, J.J.,
Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, S.Z., Chen, L., Fan, D.L.,
Weng, Q.J., Zhang, L., Lu, Y.O., Yu, S.L., Liu, X.H., Lu, T.T.,
Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X.,
Zhang, L., Fan, L.F., Chen, W., Wu, S.A. and Xue, Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences

```

